KAM for quasi-linear and fully nonlinear forced KdV

Pietro Baldi, Massimiliano Berti, Riccardo Montalto

Abstract: We prove the existence of quasi-periodic, small amplitude, solutions for quasi-linear and fully nonlinear forced perturbations of KdV equations. For Hamiltonian or reversible nonlinearities we also obtain the linear stability of the solutions. The proofs are based on a combination of different ideas and techniques: (*i*) a Nash-Moser iterative scheme in Sobolev scales. (*ii*) A regularization procedure, which conjugates the linearized operator to a differential operator with constant coefficients plus a bounded remainder. These transformations are obtained by changes of variables induced by diffeomorphisms of the torus and pseudo-differential operators. (*iii*) A reducibility KAM scheme, which completes the reduction to constant coefficients of the linearized operator, providing a sharp asymptotic expansion of the perturbed eigenvalues.

Keywords: KdV, KAM for PDEs, quasi-linear PDEs, fully nonlinear PDEs, Nash-Moser theory, quasi-periodic solutions, small divisors.

Contents

1 Introduction

One of the most challenging and open questions in KAM theory concerns its possible extension to *quasi-linear* and *fully nonlinear* PDEs, namely partial differential equations whose nonlinearities contain derivatives of the same order as the linear operator. Besides its mathematical interest, this question is also relevant in view of applications to physical real world nonlinear models, for example in fluid dynamics and elasticity.

The goal of this paper is to develop KAM theory for quasi-periodically forced KdV equations of the form

$$u_t + u_{xxx} + \varepsilon f(\omega t, x, u, u_x, u_{xx}, u_{xxx}) = 0, \quad x \in \mathbb{T} := \mathbb{R}/2\pi\mathbb{Z}. \tag{1.1}$$

First, we prove in Theorem 1.1 an existence result of quasi-periodic solutions for a large class of quasi-linear nonlinearities f. Then for Hamiltonian or reversible nonlinearities, we also prove the linear stability of the solutions, see Theorems 1.2, 1.3. Theorem 1.3 also holds for fully nonlinear perturbations. The precise meaning of stability is stated in Theorem 1.5. The key analysis is the reduction to constant coefficients of the linearized KdV equation, see Theorem 1.4. To the best of our knowledge, these are the first KAM results for quasi-linear or fully nonlinear PDEs.

Let us outline a short history of the subject. KAM and Nash-Moser theory for PDEs, which counts nowadays on a wide literature, started with the pioneering works of Kuksin [32] and Wayne [44], and was developed in the 1990s by Craig-Wayne [18], Bourgain [13], [14], Pöschel [39] (see also [34], [17] for more references). These papers concern wave and Schrödinger equations with bounded Hamiltonian nonlinearities.

The first KAM results for *unbounded* perturbations have been obtained by Kuksin [33], [34], and, then, Kappeler-Pöschel [30], for Hamiltonian, analytic perturbations of KdV. Here the highest constant coefficients linear operator is ∂_{xxx} and the nonlinearity contains one space derivative ∂_x. Their approach has been recently improved by Liu-Yuan [37] and Zhang-Gao-Yuan [45] for 1-dimensional derivative NLS (DNLS) and Benjamin-Ono equations, where the highest order constant coefficients linear operator is ∂_{xx} and the nonlinearity contains one derivative ∂_x. These methods apply to dispersive PDEs with derivatives like KdV, DNLS, the Duffing oscillator (see Bambusi-Graffi [3]), but not to derivative wave equations (DNLW) which contain first order derivatives ∂_x, ∂_t in the nonlinearity.

For DNLW, KAM theorems have been recently proved by Berti-Biasco-Procesi for both Hamiltonian [11] and reversible [12] equations. The key ingredient is an asymptotic expansion of the perturbed eigenvalues that is sufficiently accurate to impose the second order Melnikov non-resonance conditions. In this way, the scheme produces a constant coefficients normal form around the invariant torus (*reducibility*), implying the linear stability of the solution. This is achieved introducing the notion of "quasi-Töplitz" vector field, which is inspired to the concept of "quasi-Töplitz" and "Töplitz-Lipschitz" Hamiltonians, developed, respectively, in Procesi-Xu [41] and Eliasson-Kuksin [20], [21] (see also Geng-You-Xu [22], Grébert-Thomann [24], Procesi-Procesi [40]).

Existence of quasi-periodic solutions of PDEs can also be proved by imposing only the first order Melnikov conditions. This approach has been developed by Bourgain [13]-[16] extending the work of Craig-Wayne [18] for periodic solutions. It is especially convenient for PDEs in higher space dimension, because of the high multiplicity of the eigenvalues: see also the recent results by Wang [43], Berti-Bolle [8], [9] (and [4], [10], [23] for periodic solutions). This method does not provide informations about the stability of the quasi-periodic solutions, because the linearized equations have variable coefficients.

All the aforementioned results concern "semilinear" PDEs, namely equations in which the nonlinearity contains *strictly less* derivatives than the linear differential operator. For quasi-linear or fully nonlinear PDEs the perturbative effect is much stronger, and the possibility of extending KAM theory in this context is doubtful, see [30], [17], [37], because of the possible phenomenon of formation of singularities outlined in Lax [36], Klainerman and Majda [31]. For example, Kappeler-Pöschel [30] (remark 3, page 19) wrote: "*It would be interesting to obtain perturbation results which also include terms of higher order, at least in the region where the KdV approximation is valid. However, results of this type are still out of reach, if true at all*". The study of this important issue is at its infancy.

For quasi-linear and fully nonlinear PDEs, the literature concerns, so far, only existence of *periodic* solutions. We quote the classical bifurcation results of Rabinowitz [42] for fully nonlinear forced wave

equations with a small dissipation term. More recently, Baldi [1] proved existence of periodic forced vibrations for quasi-linear Kirchhoff equations. Here the quasi-linear perturbation term depends explicitly only on time. Both these results are proved via Nash-Moser methods.

For the water waves equations, which are a fully nonlinear PDE, we mention the pioneering work of Iooss-Plotnikov-Toland [27] about the existence of time periodic standing waves, and of Iooss-Plotinikov [28], [29] for 3-dimensional traveling water waves. The key idea is to use diffeomorphisms of the torus $\mathbb{T}^2$ and pseudo-differential operators, in order to conjugate the linearized operator (at an approximate solution) to a constant coefficients operator plus a sufficiently regularizing remainder. This is enough to invert the whole linearized operator by Neumann series.

Very recently Baldi [2] has further developed the techniques of [27], proving the existence of periodic solutions for fully nonlinear autonomous, reversible Benjamin-Ono equations.

These approaches do not imply the linear stability of the solutions and, unfortunately, they do not work for quasi-periodic solutions, because stronger small divisors difficulties arise, see the observation 5 below.

We finally mention that, for quasi-linear Klein-Gordon equations on spheres, Delort [19] has proved long time existence results via Birkhoff normal form methods.

In the present paper we combine different ideas and techniques. The key analysis concerns the linearized KdV operator (1.16) obtained at any step of the Nash-Moser iteration. First, we use changes of variables, like quasi-periodic time-dependent diffeomorphisms of the space variable x, a quasi-periodic reparametrization of time, multiplication operators and Fourier multipliers, in order to reduce the linearized operator to constant coefficients up to a bounded remainder, see (1.24). These transformations, which are inspired to [2], [27], are very different from the usual KAM transformations. Then, we perform a quadratic KAM reducibility scheme *à la* Eliasson-Kuksin, which completely diagonalizes the linearized operator. For reversible or Hamiltonian KdV perturbations we get that the eigenvalues of this diagonal operator are purely imaginary, i.e. we prove the linear stability. In section 1.2 we present the main ideas of proof.

We remark that the present approach could be also applied to quasi-linear and fully nonlinear perturbations of dispersive PDEs like 1-dimensional NLS and Benjamin-Ono equations (but not to the wave equation, which is not dispersive). For definiteness, we have developed all the computations in KdV case.

In the next subsection we state precisely our KAM results. In order to highlight the main ideas, we consider the simplest setting of nonlinear perturbations of the Airy-KdV operator $\partial_t + \partial_{xxx}$ and we look for small amplitude solutions.

1.1 Main results

We consider problem (1.1) where $\varepsilon > 0$ is a small parameter, the nonlinearity is quasi-periodic in time with diophantine frequency vector

$$\omega = \lambda\bar{\omega} \in \mathbb{R}^\nu, \quad \lambda \in \Lambda := \left[\frac{1}{2}, \frac{3}{2}\right], \quad |\bar{\omega} \cdot l| \geq \frac{3\gamma_0}{|l|^{\tau_0}} \quad \forall l \in \mathbb{Z}^\nu \setminus \{0\}, \tag{1.2}$$

and $f(\varphi, x, z)$, $\varphi \in \mathbb{T}^\nu$, $z := (z_0, z_1, z_2, z_3) \in \mathbb{R}^4$, is a finitely many times differentiable function, namely

$$f \in C^q(\mathbb{T}^\nu \times \mathbb{T} \times \mathbb{R}^4; \mathbb{R}) \tag{1.3}$$

for some $q \in \mathbb{N}$ large enough. For simplicity we fix in (1.2) the diophantine exponent $\tau_0 := \nu$. The only "external" parameter in (1.1) is λ, which is the length of the frequency vector (this corresponds to a time scaling).

We consider the following questions:

- *For ε small enough, do there exist quasi-periodic solutions of (1.1) for positive measure sets of $\lambda \in \Lambda$?*
- *Are these solutions linearly stable?*

Clearly, if $f(\varphi, x, 0)$ is not identically zero, then $u = 0$ is not a solution of (1.1) for $\varepsilon \neq 0$. Thus we look for non-trivial $(2\pi)^{\nu+1}$-periodic solutions $u(\varphi, x)$ of

$$\omega \cdot \partial_\varphi u + u_{xxx} + \varepsilon f(\varphi, x, u, u_x, u_{xx}, u_{xxx}) = 0 \tag{1.4}$$

in the Sobolev space

$$H^s := H^s(\mathbb{T}^\nu \times \mathbb{T}; \mathbb{R}) \tag{1.5}$$
$$:= \left\{ u(\varphi, x) = \sum_{(l,j) \in \mathbb{Z}^\nu \times \mathbb{Z}} u_{l,j}\, e^{\mathrm{i}(l\cdot\varphi + jx)} \in \mathbb{R}, \quad \bar{u}_{l,j} = u_{-l,-j}, \quad \|u\|_s^2 := \sum_{(l,j) \in \mathbb{Z}^\nu \times \mathbb{Z}} \langle l,j\rangle^{2s} |u_{l,j}|^2 < \infty \right\}$$

where

$$\langle l, j \rangle := \max\{1, |l|, |j|\}.$$

From now on, we fix $\mathfrak{s}_0 := (\nu + 2)/2 > (\nu + 1)/2$, so that for all $s \geq \mathfrak{s}_0$ the Sobolev space H^s is a Banach algebra, and it is continuously embedded $H^s(\mathbb{T}^{\nu+1}) \hookrightarrow C(\mathbb{T}^{\nu+1})$.

We need some assumptions on the nonlinearity. We consider *fully nonlinear* perturbations satisfying

- TYPE (F)

$$\partial_{z_2} f = 0, \tag{1.6}$$

namely f is independent of u_{xx}. Otherwise, we require that

- TYPE (Q)

$$\partial_{z_3 z_3}^2 f = 0, \quad \partial_{z_2} f = \alpha(\varphi)\Big(\partial_{z_3 x}^2 f + z_1 \partial_{z_3 z_0}^2 f + z_2 \partial_{z_3 z_1}^2 f + z_3 \partial_{z_3 z_2}^2 f\Big) \tag{1.7}$$

for some function $\alpha(\varphi)$ (independent on x).

If (Q) holds, then the nonlinearity f depends linearly on u_{xxx}, namely equation (1.1) is *quasi-linear*. We note that the Hamiltonian nonlinearities, see (1.11), are a particular case of those satisfying (Q), see remark 3.2. In comment 3 after Theorem 1.5 we explain the reason for assuming either condition (F) or (Q).

The following theorem is an existence result of quasi-periodic solutions for quasi-linear KdV equations.

Theorem 1.1. (Existence) *There exist $s := s(\nu) > 0$, $q := q(\nu) \in \mathbb{N}$, such that:*
For every quasi-linear nonlinearity $f \in C^q$ of the form

$$f = \partial_x\big(g(\omega t, x, u, u_x, u_{xx})\big) \tag{1.8}$$

satisfying the (Q)-condition (1.7), for all $\varepsilon \in (0, \varepsilon_0)$, where $\varepsilon_0 := \varepsilon_0(f, \nu)$ is small enough, there exists a Cantor set $\mathcal{C}_\varepsilon \subset \Lambda$ of asymptotically full Lebesgue measure, i.e.

$$|\mathcal{C}_\varepsilon| \to 1 \quad as \quad \varepsilon \to 0, \tag{1.9}$$

such that, $\forall \lambda \in \mathcal{C}_\varepsilon$ the perturbed KdV equation (1.4) has a solution $u(\varepsilon, \lambda) \in H^s$ with $\|u(\varepsilon, \lambda)\|_s \to 0$ as $\varepsilon \to 0$.

We may ensure the *linear stability* of the solutions requiring further conditions on the nonlinearity, see Theorem 1.5 for the precise statement. The first case is that of *Hamiltonian* KdV equations

$$u_t = \partial_x \nabla_{L^2} H(t, x, u, u_x), \quad H(t, x, u, u_x) := \int_\mathbb{T} \frac{u_x^2}{2} + \varepsilon F(\omega t, x, u, u_x)\, dx \tag{1.10}$$

which have the form (1.1), (1.8) with

$$f(\varphi, x, u, u_x, u_{xx}, u_{xxx}) = -\partial_x\big\{(\partial_{z_0} F)(\varphi, x, u, u_x)\big\} + \partial_{xx}\big\{(\partial_{z_1} F)(\varphi, x, u, u_x)\big\}. \tag{1.11}$$

The phase space of (1.10) is

$$H_0^1(\mathbb{T}) := \left\{ u(x) \in H^1(\mathbb{T}, \mathbb{R}) \; : \; \int_\mathbb{T} u(x)\, dx = 0 \right\}$$

endowed with the non-degenerate symplectic form

$$\Omega(u, v) := \int_\mathbb{T} (\partial_x^{-1} u) v\, dx, \quad \forall u, v \in H_0^1(\mathbb{T}), \tag{1.12}$$

where $\partial_x^{-1} u$ is the periodic primitive of u with zero average, see (3.19). As proved in remark 3.2, the Hamiltonian nonlinearity f in (1.11) satisfies also the (Q)-condition (1.7). As a consequence, Theorem 1.1 implies the existence of quasi-periodic solutions of (1.10). In addition, we also prove their linear stability.

Theorem 1.2. (Hamiltonian KdV) *For all Hamiltonian quasi-linear KdV equations* (1.10) *the quasi-periodic solution* $u(\varepsilon, \lambda)$ *found in Theorem 1.1 is* LINEARLY STABLE *(see Theorem 1.5).*

The stability of the quasi-periodic solutions also follows by the *reversibility* condition

$$f(-\varphi, -x, z_0, -z_1, z_2, -z_3) = -f(\varphi, x, z_0, z_1, z_2, z_3).$$ (1.13)

Actually (1.13) implies that the infinite-dimensional non-autonomous dynamical system

$$u_t = V(t, u), \quad V(t, u) := -u_{xxx} - \varepsilon f(\omega t, x, u, u_x, u_{xx}, u_{xxx})$$

is reversible with respect to the involution

$$S : u(x) \to u(-x), \quad S^2 = I,$$

namely

$$-SV(-t, u) = V(t, Su).$$

In this case it is natural to look for "reversible" solutions of (1.4), that is

$$u(\varphi, x) = u(-\varphi, -x).$$ (1.14)

Theorem 1.3. (Reversible KdV) *There exist* $s := s(\nu) > 0$, $q := q(\nu) \in \mathbb{N}$, *such that:*
For every nonlinearity $f \in C^q$ *that satisfies*

   (i) *the reversibility condition* (1.13),

and

   (ii) *either the (F)-condition* (1.6) *or the (Q)-condition* (1.7),

for all $\varepsilon \in (0, \varepsilon_0)$, *where* $\varepsilon_0 := \varepsilon_0(f, \nu)$ *is small enough, there exists a Cantor set* $\mathcal{C}_\varepsilon \subset \Lambda$ *with Lebesgue measure satisfying* (1.9), *such that for all* $\lambda \in \mathcal{C}_\varepsilon$ *the perturbed KdV equation* (1.4) *has a solution* $u(\varepsilon, \lambda) \in H^s$ *that satisfies* (1.14), *with* $\|u(\varepsilon, \lambda)\|_s \to 0$ *as* $\varepsilon \to 0$. *In addition,* $u(\varepsilon, \lambda)$ *is* LINEARLY STABLE.

Let us make some comments on the results.

1. The previous theorems (in particular the Hamiltonian Theorem 1.2) give a positive answer to the question that was posed by Kappeler-Pöschel [30], page 19, Remark 3, about the possibility of KAM type results for quasi-linear perturbations of KdV.

2. In Theorem 1.1 we do not have informations about the linear stability of the solutions because the nonlinearity f has no special structure and it may happen that some eigenvalues of the linearized operator have non zero real part (partially hyperbolic tori). We remark that, in any case, we may compute the eigenvalues (i.e. Lyapunov exponents) of the linearized operator with any order of accuracy. With further conditions on the nonlinearity—like reversibility or in the Hamiltonian case—the eigenvalues are purely imaginary, and the torus is linearly stable. The present situation is very different with respect to [18], [13]-[16], [8]-[9] and also [27]-[29], [2], where the lack of stability informations is due to the fact that the linearized equation has variable coefficients, and it is not reduced as in Theorem 1.4 below.

3. One cannot expect the existence of quasi-periodic solutions of (1.4) for *any* perturbation f. Actually, if $f = m \neq 0$ is a constant, then, integrating (1.4) in (φ, x) we find the contradiction $\varepsilon m = 0$. This is a consequence of the fact that

$$\text{Ker}(\omega \cdot \partial_\varphi + \partial_{xxx}) = \mathbb{R}$$ (1.15)

is non trivial. Both the condition (1.8) (which is satisfied by the Hamiltonian nonlinearities) and the reversibility condition (1.13) allow to overcome this obstruction, working in a space of functions with zero average. The degeneracy (1.15) also reflects in the fact that the solutions of (1.4) appear as a 1-dimensional family $c + u_c(\varepsilon, \lambda)$ parametrized by the "average" $c \in \mathbb{R}$. We could also avoid this degeneracy by adding a "mass" term $+mu$ in (1.1), but it does not seem to have physical meaning.

4. In Theorem 1.1 we have not considered the case in which f is fully nonlinear and satisfies condition (F) in (1.6), because any nonlinearity of the form (1.8) is automatically quasi-linear (and so the first condition in (1.7) holds) and (1.6) trivially implies the second condition in (1.7) with $\alpha(\varphi) = 0$.

5. The solutions $u \in H^s$ have the same regularity in both variables (φ, x). This functional setting is convenient when using changes of variables that mix the time and space variables, like the composition operators $\mathcal{A}$, $\mathcal{T}$ in sections 3.1, 3.4,

6. In the Hamiltonian case (1.10), the nonlinearity f in (1.11) satisfies the reversibility condition (1.13) if and only if $F(-\varphi, -x, z_0, -z_1) = F(\varphi, x, z_0, z_1)$.

Theorems 1.1-1.3 are based on a Nash-Moser iterative scheme. An essential ingredient in the proof—which also implies the linear stability of the quasi-periodic solutions—is the *reducibility* of the linear operator

$$\mathcal{L} := \mathcal{L}(u) = \omega \cdot \partial_\varphi + (1 + a_3(\varphi, x))\partial_{xxx} + a_2(\varphi, x)\partial_{xx} + a_1(\varphi, x)\partial_x + a_0(\varphi, x) \qquad (1.16)$$

obtained linearizing (1.4) at any approximate (or exact) solution u, namely the coefficients $a_i(\varphi, x)$ are defined in (3.2). Let $H_x^s := H^s(\mathbb{T})$ denote the usual Sobolev spaces of functions of $x \in \mathbb{T}$ only (phase space).

Theorem 1.4. (Reducibility) *There exist $\bar{\sigma} > 0$, $q \in \mathbb{N}$, depending on ν, such that:*

For every nonlinearity $f \in C^q$ that satisfies the hypotheses of Theorems 1.1 or 1.3, for all $\varepsilon \in (0, \varepsilon_0)$, where $\varepsilon_0 := \varepsilon_0(f, \nu)$ is small enough, for all u in the ball $\|u\|_{\mathfrak{s}_0 + \bar{\sigma}} \leq 1$, there exists a Cantor like set $\Lambda_\infty(u) \subset \Lambda$ such that, for all $\lambda \in \Lambda_\infty(u)$:

i) for all $s \in (\mathfrak{s}_0, q - \bar{\sigma})$, if $\|u\|_{s+\bar{\sigma}} < +\infty$ then there exist linear invertible bounded operators W_1, W_2 : $H^s(\mathbb{T}^{\nu+1}) \to H^s(\mathbb{T}^{\nu+1})$ with bounded inverse, that semi-conjugate the linear operator $\mathcal{L}(u)$ in (1.16) to the diagonal operator $\mathcal{L}_\infty$, namely

$$\mathcal{L}(u) = W_1 \mathcal{L}_\infty W_2^{-1}, \quad \mathcal{L}_\infty := \omega \cdot \partial_\varphi + \mathcal{D}_\infty \qquad (1.17)$$

where

$$\mathcal{D}_\infty := \operatorname{diag}_{j \in \mathbb{Z}}\{\mu_j\}, \quad \mu_j := \mathrm{i}(-m_3 j^3 + m_1 j) + r_j, \quad m_3, m_1 \in \mathbb{R}, \quad \sup_j |r_j| \leq C\varepsilon. \qquad (1.18)$$

ii) For each $\varphi \in \mathbb{T}^\nu$ the operators W_i are also bounded linear bijections of the phase space (see notation (2.18))

$$W_i(\varphi), W_i^{-1}(\varphi) : H_x^s \to H_x^s, \quad i = 1, 2.$$

A curve $h(t) = h(t, \cdot) \in H_x^s$ is a solution of the quasi-periodically forced linear KdV equation

$$\partial_t h + (1 + a_3(\omega t, x))\partial_{xxx} h + a_2(\omega t, x)\partial_{xx} h + a_1(\omega t, x)\partial_x h + a_0(\omega t, x)h = 0 \qquad (1.19)$$

if and only if the transformed curve

$$v(t) := v(t, \cdot) := W_2^{-1}(\omega t)[h(t)] \in H_x^s$$

is a solution of the constant coefficients dynamical system

$$\partial_t v + \mathcal{D}_\infty v = 0, \quad \dot{v}_j = -\mu_j v_j, \quad \forall j \in \mathbb{Z}. \qquad (1.20)$$

In the reversible or Hamiltonian case all the $\mu_j \in \mathrm{i}\mathbb{R}$ are purely imaginary.

The exponents μ_j can be effectively computed. All the solutions of (1.20) are

$$v(t) = \sum_{j \in \mathbb{Z}} v_j(t)e^{\mathrm{i}jx}, \quad v_j(t) = e^{-\mu_j t}v_j(0).$$

If the μ_j are purely imaginary – as in the reversible or the Hamiltonian cases – all the solutions of (1.20) are almost periodic in time (in general) and the Sobolev norm

$$\|v(t)\|_{H_x^s} = \Big(\sum_{j \in \mathbb{Z}} |v_j(t)|^2 \langle j \rangle^{2s}\Big)^{1/2} = \Big(\sum_{j \in \mathbb{Z}} |v_j(0)|^2 \langle j \rangle^{2s}\Big)^{1/2} = \|v(0)\|_{H_x^s} \qquad (1.21)$$

is constant in time. As a consequence we have:

Theorem 1.5. (Linear stability) *Assume the hypothesis of Theorem 1.4 and, in addition, that f is Hamiltonian (see (1.11)) or it satisfies the reversibility condition (1.13). Then, $\forall s \in (\mathfrak{s}_0, q - \bar{\sigma} - \mathfrak{s}_0)$, $\|u\|_{s+\mathfrak{s}_0+\bar{\sigma}} < +\infty$, there exists $K_0 > 0$ such that for all $\lambda \in \Lambda_\infty(u)$, $\varepsilon \in (0, \varepsilon_0)$, all the solutions of (1.19) satisfy*

$$\|h(t)\|_{H_x^s} \le K_0 \|h(0)\|_{H_x^s} \tag{1.22}$$

and, for some $\mathtt{a} \in (0, 1)$,

$$\|h(0)\|_{H_x^s} - \varepsilon^{\mathtt{a}} K_0 \|h(0)\|_{H_x^{s+1}} \le \|h(t)\|_{H_x^s} \le \|h(0)\|_{H_x^s} + \varepsilon^{\mathtt{a}} K_0 \|h(0)\|_{H_x^{s+1}}. \tag{1.23}$$

Theorems 1.1-1.5 are proved in section 5.1 collecting all the informations of sections 2-5.

1.2 Ideas of proof

The proof of Theorems 1.1-1.3 is based on a Nash-Moser iterative scheme in the scale of Sobolev spaces H^s. The main issue concerns the invertibility of the linearized KdV operator $\mathcal{L}$ in (1.16), at each step of the iteration, and the proof of the tame estimates (5.7) for its right inverse. This information is obtained in Theorem 4.3 by conjugating $\mathcal{L}$ to constant coefficients. This is also the key which implies the stability results for the Hamiltonian and reversible nonlinearities, see Theorems 1.4-1.5.

We now explain the main ideas of the reducibility scheme. The term of $\mathcal{L}$ that produces the strongest perturbative effects to the spectrum (and eigenfunctions) is $a_3(\varphi, x)\partial_{xxx}$, and, then $a_2(\varphi, x)\partial_{xx}$. The usual KAM transformations are not able to deal with these terms because they are "too close" to the identity. Our strategy is the following. First, we conjugate the operator $\mathcal{L}$ in (1.16) to a constant coefficients third order differential operator plus a zero order remainder

$$\mathcal{L}_5 = \omega \cdot \partial_\varphi + m_3 \partial_{xxx} + m_1 \partial_x + \mathcal{R}_0, \quad m_3 = 1 + O(\varepsilon), \ m_1 = O(\varepsilon), \ m_1, m_3 \in \mathbb{R}, \tag{1.24}$$

(see (3.55)), via changes of variables induced by diffeomorphisms of the torus, reparametrization of time, and pseudo-differential operators. This is the goal of section 3. All these transformations could be composed into one map, but we find it more convenient to split the regularization procedure into separate steps (sections 3.1-3.5), both to highlight the basic ideas, and, especially, in order to derive estimates on the coefficients, section 3.6. Let us make some comments on this procedure.

1. In order to eliminate the space variable dependence of the highest order perturbation $a_3(\varphi, x)\partial_{xxx}$ (see (3.20)) we use, in section 3.1, φ-dependent changes of variables like

   $$(\mathcal{A}h)(\varphi, x) := h(\varphi, x + \beta(\varphi, x)).$$

   These transformations converge pointwise to the identity if $\beta \to 0$ but not in operatorial norm. If β is odd, $\mathcal{A}$ preserves the reversible structure, see remark 3.4. On the other hand for the Hamiltonian KdV (1.10) we use the modified transformation

   $$(\mathcal{A}h)(\varphi, x) := (1 + \beta_x(\varphi, x)) h(\varphi, x + \beta(\varphi, x)) = \frac{d}{dx}\big\{(\partial_x^{-1}h)(\varphi, x + \beta(\varphi, x))\big\} \tag{1.25}$$

   for all $h(\varphi, \cdot) \in H_0^1(\mathbb{T})$. This map is canonical, for each $\varphi \in \mathbb{T}^\nu$, with respect to the KdV-symplectic form (1.12), see remark 3.3. Thus (1.25) preserves the Hamiltonian structure and also eliminates the term of order ∂_{xx}, see remark 3.5.

2. In the second step of section 3.2 we eliminate the time dependence of the coefficients of the highest order spatial derivative operator ∂_{xxx} by a quasi-periodic time re-parametrization. This procedure preserves the reversible and the Hamiltonian structure, see remark 3.6 and 3.7.

3. Assumptions (Q) (see (1.7)) or (F) (see (1.6)) allow to eliminate terms like $a(\varphi, x)\partial_{xx}$ along this reduction procedure, see (3.41). This is possible, by a conjugation with multiplication operators (see (3.34)), if (see (3.40))

   $$\int_{\mathbb{T}} \frac{a_2(\varphi, x)}{1 + a_3(\varphi, x)}\, dx = 0. \tag{1.26}$$

If (F) holds, then the coefficient $a_2(\varphi, x) = 0$ and (1.26) is satisfied. If (Q) holds, then an easy computation shows that $a_2(\varphi, x) = \alpha(\varphi) \, \partial_x a_3(\varphi, x)$ (using the explicit expression of the coefficients in (3.2)), and so

$$\int_{\mathbb{T}} \frac{a_2(\varphi, x)}{1 + a_3(\varphi, x)} \, dx = \int_{\mathbb{T}} \alpha(\varphi) \, \partial_x \big(\log[1 + a_3(\varphi, x)] \big) \, dx = 0 \, .$$

In both cases (Q) and (F), condition (1.26) is satisfied.

In the Hamiltonian case there is no need of this step because the symplectic transformation (1.25) also eliminates the term of order ∂_{xx}, see remark 3.7.

We note that without assumptions (Q) or (F) we may always reduce $\mathcal{L}$ to a time dependent operator with $a(\varphi)\partial_{xx}$. If $a(\varphi)$ were a constant, then this term would even simplify the analysis, killing the small divisors. The pathological situation that we want to eliminate assuming (Q) or (F) is when $a(\varphi)$ changes sign. In such a case, this term acts as a friction when $a(\varphi) < 0$ and as an amplifier when $a(\varphi) > 0$.

4. In sections 3.4-3.5, we are finally able to conjugate the linear operator to another one with a coefficient in front of ∂_x which is constant, i.e. obtaining (1.24). In this step we use a transformation of the form $I + w(\varphi, x)\partial_x^{-1}$, see (3.49). In the Hamiltonian case we use the symplectic map $e^{\pi_0 w(\varphi, x)\partial_x^{-1}}$, see remark 3.13.

5. We can iterate the regularization procedure at any *finite* order $k = 0, 1, \ldots$, conjugating $\mathcal{L}$ to an operator of the form $\mathfrak{D} + \mathcal{R}$, where

$$\mathfrak{D} = \omega \cdot \partial_\varphi + \mathcal{D}, \quad \mathcal{D} = m_3 \partial_x^3 + m_1 \partial_x + \ldots + m_{-k} \partial_x^{-k}, \quad m_i \in \mathbb{R},$$

has constant coefficients, and the rest $\mathcal{R}$ is arbitrarily regularizing in space, namely

$$\partial_x^k \circ \mathcal{R} = \text{bounded} \, . \tag{1.27}$$

However, one cannot iterate this regularization infinitely many times, because it is not a quadratic scheme, and therefore, because of the small divisors, it does not converge. This regularization procedure is sufficient to prove the invertibility of $\mathcal{L}$, giving tame estimates for the inverse, in the periodic case, but it does not work for quasi-periodic solutions. The reason is the following. In order to use Neumann series, one needs that $\mathfrak{D}^{-1}\mathcal{R} = (\mathfrak{D}^{-1}\partial_x^{-k})(\partial_x^k \mathcal{R})$ is bounded, namely, in view of (1.27), that $\mathfrak{D}^{-1}\partial_x^{-k}$ is bounded. In the region where the eigenvalues $(\mathrm{i}\omega \cdot l + \mathcal{D}_j)$ of $\mathfrak{D}$ are small, space and time derivatives are related, $|\omega \cdot l| \sim |j|^3$, where l is the Fourier index of time, j is that of space, and $\mathcal{D}_j = -\mathrm{i}m_3 j^3 + \mathrm{i}m_1 j + \ldots$ are the eigenvalues of $\mathcal{D}$. Imposing the first order Melnikov conditions $|\mathrm{i}\omega \cdot l + \mathcal{D}_j| > \gamma |l|^{-\tau}$, in that region, $(\mathfrak{D}^{-1}\partial_x^{-k})$ has eigenvalues

$$\left| \frac{1}{(\mathrm{i}\omega \cdot l + \mathcal{D}_j)j^k} \right| < \frac{|l|^\tau}{\gamma |j|^k} < \frac{C|l|^\tau}{|\omega \cdot l|^{k/3}} \, .$$

In the periodic case, $\omega \in \mathbb{R}$, $l \in \mathbb{Z}$, $|\omega \cdot l| = |\omega||l|$, and this determines the order of regularization that is required by the procedure: $k \geq 3\tau$. In the quasi-periodic case, instead, $|l|$ is not controlled by $|\omega \cdot l|$, and the argument fails.

Once (1.24) has been obtained, we implement a quadratic reducibility KAM scheme to diagonalize $\mathcal{L}_5$, namely to conjugate $\mathcal{L}_5$ to the diagonal operator $\mathcal{L}_\infty$ in (1.17). Since we work with finite regularity, we perform a Nash-Moser smoothing regularization (time-Fourier truncation). We use standard KAM transformations, in order to decrease, quadratically at each step, the size of the perturbation $\mathcal{R}$, see section 4.1.1. This iterative scheme converges (Theorem 4.2) because the initial remainder $\mathcal{R}_0$ is a bounded operator (of the space variable x), and this property is preserved along the iteration. This is the reason for performing the regularization procedure of sections 3.1-3.5. We manage to impose the second order Melnikov non-resonance conditions (4.17), which are required by the reducibility scheme, thanks to the good control of the eigenvalues $\mu_j = -\mathrm{i}m_3(\varepsilon, \lambda)j^3 + \mathrm{i}m_1(\varepsilon, \lambda)j + r_j(\varepsilon, \lambda)$, where $\sup_j |r_j(\varepsilon, \lambda)| = O(\varepsilon)$.

Note that the eigenvalues μ_j could be not purely imaginary, i.e. r_j could have a non-zero real part which depends on the nonlinearity (unlike the reversible or Hamiltonian case, where $r_j \in i\mathbb{R}$). In such a case, the invariant torus could be (partially) hyperbolic. Since we do not control the real part of r_j (i.e. the hyperbolicity may vanish), we perform the measure estimates proving the diophantine lower bounds of the imaginary part of the small divisors.

The final comment concerns the dynamical consequences of Theorem 1.4-ii). All the above transformations (both the changes of variables of sections 3.1-3.5 as well as the KAM matrices of the reducibility scheme) are time-dependent quasi-periodic maps of the phase space (of functions of x only), see section 2.2. It is thanks to this "Töplitz-in-time" structure that the linear KdV equation (1.19) is transformed into the dynamical system (1.20). Note that in [27] (and also [16], [8],[9]) the analogous transformations have not this Töplitz-in-time structure and stability informations are not obtained.

2 Functional setting

For a function $f : \Lambda_o \to E$, $\lambda \mapsto f(\lambda)$, where $(E, \| \ \|_E)$ is a Banach space and Λ_o is a subset of $\mathbb{R}$, we define the sup-norm and the Lipschitz semi-norm

$$\|f\|_E^{\sup} := \|f\|_{E,\Lambda_o}^{\sup} := \sup_{\lambda \in \Lambda_o} \|f(\lambda)\|_E , \quad \|f\|_E^{\mathrm{lip}} := \|f\|_{E,\Lambda_o}^{\mathrm{lip}} := \sup_{\substack{\lambda_1, \lambda_2 \in \Lambda_o \\ \lambda_1 \neq \lambda_2}} \frac{\|f(\lambda_1) - f(\lambda_2)\|_E}{|\lambda_1 - \lambda_2|} , \qquad (2.1)$$

and, for $\gamma > 0$, the Lipschitz norm

$$\|f\|_E^{\mathrm{Lip}(\gamma)} := \|f\|_{E,\Lambda_o}^{\mathrm{Lip}(\gamma)} := \|f\|_E^{\sup} + \gamma \|f\|_E^{\mathrm{lip}} . \qquad (2.2)$$

If $E = H^s$ we simply denote $\|f\|_{H^s}^{\mathrm{Lip}(\gamma)} := \|f\|_s^{\mathrm{Lip}(\gamma)}$.

As a notation, we write

$$a \leq_s b \quad \Longleftrightarrow \quad a \leq C(s)b$$

for some constant $C(s)$. For $s = \mathfrak{s}_0 := (\nu + 2)/2$ we only write $a \lessdot b$. More in general the notation $a \lessdot b$ means $a \leq Cb$ where the constant C may depend on the data of the problem, namely the nonlinearity f, the number ν of frequencies, the diophantine vector $\bar{\omega}$, the diophantine exponent $\tau > 0$ in the non-resonance conditions in (4.6). Also the small constants δ in the sequel depend on the data of the problem.

2.1 Matrices with off-diagonal decay

Let $b \in \mathbb{N}$ and consider the exponential basis $\{e_i : i \in \mathbb{Z}^b\}$ of $L^2(\mathbb{T}^b)$, so that $L^2(\mathbb{T}^b)$ is the vector space $\{u = \sum u_i e_i, \sum |u_i|^2 < \infty\}$. Any linear operator $A : L^2(\mathbb{T}^b) \to L^2(\mathbb{T}^b)$ can be represented by the infinite dimensional matrix

$$(A_i^{i'})_{i,i' \in \mathbb{Z}^b}, \quad A_i^{i'} := (Ae_{i'}, e_i)_{L^2(\mathbb{T}^b)}, \quad Au = \sum_{i,i'} A_i^{i'} u_{i'} e_i.$$

We now define the s-norm (introduced in [8]) of an infinite dimensional matrix.

Definition 2.1. *The s-decay norm of an infinite dimensional matrix $A := (A_{i_1}^{i_2})_{i_1, i_2 \in \mathbb{Z}^b}$ is*

$$|A|_s^2 := \sum_{i \in \mathbb{Z}^b} \langle i \rangle^{2s} \left(\sup_{i_1 - i_2 = i} |A_{i_1}^{i_2}| \right)^2 . \qquad (2.3)$$

For parameter dependent matrices $A := A(\lambda)$, $\lambda \in \Lambda_o \subseteq \mathbb{R}$, the definitions (2.1) and (2.2) become

$$|A|_s^{\sup} := \sup_{\lambda \in \Lambda_o} |A(\lambda)|_s , \quad |A|_s^{\mathrm{lip}} := \sup_{\lambda_1 \neq \lambda_2} \frac{|A(\lambda_1) - A(\lambda_2)|_s}{|\lambda_1 - \lambda_2|} , \quad |A|_s^{\mathrm{Lip}(\gamma)} := |A|_s^{\sup} + \gamma |A|_s^{\mathrm{lip}} .$$

Clearly, the matrix decay norm (2.3) is increasing with respect to the index s, namely

$$|A|_s \leq |A|_{s'}, \quad \forall s < s'.$$

The s-norm is designed to estimate the polynomial off-diagonal decay of matrices, actually it implies

$$|A_{i_1}^{i_2}| \leq \frac{|A|_s}{\langle i_1 - i_2 \rangle^s}, \quad \forall i_1, i_2 \in \mathbb{Z}^b,$$

and, on the diagonal elements,

$$|A_i^i| \leq |A|_0, \quad |A_i^i|^{\mathrm{lip}} \leq |A|_0^{\mathrm{lip}}. \tag{2.4}$$

We now list some properties of the matrix decay norm proved in [8].

Lemma 2.1. (Multiplication operator) *Let $p = \sum_i p_i e_i \in H^s(\mathbb{T}^b)$. The multiplication operator $h \mapsto ph$ is represented by the Töplitz matrix $T_i^{i'} = p_{i-i'}$ and*

$$|T|_s = \|p\|_s. \tag{2.5}$$

Moreover, if $p = p(\lambda)$ is a Lipschitz family of functions,

$$|T|_s^{\mathrm{Lip}(\gamma)} = \|p\|_s^{\mathrm{Lip}(\gamma)}. \tag{2.6}$$

The s-norm satisfies classical algebra and interpolation inequalities.

Lemma 2.2. (Interpolation) *For all $s \geq s_0 > b/2$ there are $C(s) \geq C(s_0) \geq 1$ such that*

$$|AB|_s \leq C(s)|A|_{s_0}|B|_s + C(s_0)|A|_s|B|_{s_0}. \tag{2.7}$$

In particular, the algebra property holds

$$|AB|_s \leq C(s)|A|_s|B|_s. \tag{2.8}$$

If $A = A(\lambda)$ and $B = B(\lambda)$ depend in a Lipschitz way on the parameter $\lambda \in \Lambda_o \subset \mathbb{R}$, then

$$|AB|_s^{\mathrm{Lip}(\gamma)} \leq C(s)|A|_s^{\mathrm{Lip}(\gamma)}|B|_s^{\mathrm{Lip}(\gamma)}, \tag{2.9}$$

$$|AB|_s^{\mathrm{Lip}(\gamma)} \leq C(s)|A|_s^{\mathrm{Lip}(\gamma)}|B|_{s_0}^{\mathrm{Lip}(\gamma)} + C(s_0)|A|_{s_0}^{\mathrm{Lip}(\gamma)}|B|_s^{\mathrm{Lip}(\gamma)}. \tag{2.10}$$

For all $n \geq 1$, using (2.8) with $s = s_0$, we get

$$|A^n|_{s_0} \leq [C(s_0)]^{n-1}|A|_{s_0}^n \qquad \text{and} \qquad |A^n|_s \leq n[C(s_0)|A|_{s_0}]^{n-1}C(s)|A|_s, \ \forall s \geq s_0. \tag{2.11}$$

Moreover (2.10) implies that (2.11) also holds for Lipschitz norms $|\ |_s^{\mathrm{Lip}(\gamma)}$.

The s-decay norm controls the Sobolev norm, also for Lipschitz families:

$$\|Ah\|_s \leq C(s)\big(|A|_{s_0}\|h\|_s + |A|_s\|h\|_{s_0}\big), \quad \|Ah\|_s^{\mathrm{Lip}(\gamma)} \leq C(s)\big(|A|_{s_0}^{\mathrm{Lip}(\gamma)}\|h\|_s^{\mathrm{Lip}(\gamma)} + |A|_s^{\mathrm{Lip}(\gamma)}\|h\|_{s_0}^{\mathrm{Lip}(\gamma)}\big). \tag{2.12}$$

Lemma 2.3. *Let $\Phi = I + \Psi$ with $\Psi := \Psi(\lambda)$, depending in a Lipschitz way on the parameter $\lambda \in \Lambda_o \subset \mathbb{R}$, such that $C(s_0)|\Psi|_{s_0}^{\mathrm{Lip}(\gamma)} \leq 1/2$. Then Φ is invertible and, for all $s \geq s_0 > b/2$,*

$$|\Phi^{-1} - I|_s \leq C(s)|\Psi|_s, \quad |\Phi^{-1}|_{s_0}^{\mathrm{Lip}(\gamma)} \leq 2, \quad |\Phi^{-1} - I|_s^{\mathrm{Lip}(\gamma)} \leq C(s)|\Psi|_s^{\mathrm{Lip}(\gamma)}. \tag{2.13}$$

If $\Phi_i = I + \Psi_i$, $i = 1, 2$, satisfy $C(s_0)|\Psi_i|_{s_0}^{\mathrm{Lip}(\gamma)} \leq 1/2$, then

$$|\Phi_2^{-1} - \Phi_1^{-1}|_s \leq C(s)\big(|\Psi_2 - \Psi_1|_s + \big(|\Psi_1|_s + |\Psi_2|_s\big)|\Psi_2 - \Psi_1|_{s_0}\big). \tag{2.14}$$

PROOF. Estimates (2.13) follow by Neumann series and (2.11). To prove (2.14), observe that

$$\Phi_2^{-1} - \Phi_1^{-1} = \Phi_1^{-1}(\Phi_1 - \Phi_2)\Phi_2^{-1} = \Phi_1^{-1}(\Psi_1 - \Psi_2)\Phi_2^{-1}$$

and use (2.7), (2.13). ∎

2.1.1 Töplitz-in-time matrices

Let now $b := \nu + 1$ and

$$e_i(\varphi, x) := e^{\mathrm{i}(l \cdot \varphi + jx)}, \quad i := (l, j) \in \mathbb{Z}^b, \quad l \in \mathbb{Z}^\nu, \quad j \in \mathbb{Z}.$$

An important sub-algebra of matrices is formed by the matrices Töplitz in time defined by

$$A^{(l_2, j_2)}_{(l_1, j_1)} := A^{j_2}_{j_1}(l_1 - l_2), \tag{2.15}$$

whose decay norm (2.3) is

$$|A|^2_s = \sum_{j \in \mathbb{Z}, l \in \mathbb{Z}^\nu} \sup_{j_1 - j_2 = j} |A^{j_2}_{j_1}(l)|^2 \langle l, j \rangle^{2s}. \tag{2.16}$$

These matrices are identified with the φ-dependent family of operators

$$A(\varphi) := \left(A^{j_2}_{j_1}(\varphi) \right)_{j_1, j_2 \in \mathbb{Z}}, \quad A^{j_2}_{j_1}(\varphi) := \sum_{l \in \mathbb{Z}^\nu} A^{j_2}_{j_1}(l) e^{\mathrm{i}l \cdot \varphi} \tag{2.17}$$

which act on functions of the x-variable as

$$A(\varphi) : h(x) = \sum_{j \in \mathbb{Z}} h_j e^{\mathrm{i}jx} \mapsto A(\varphi)h(x) = \sum_{j_1, j_2 \in \mathbb{Z}} A^{j_2}_{j_1}(\varphi) h_{j_2} e^{\mathrm{i}j_1 x}. \tag{2.18}$$

We still denote by $|A(\varphi)|_s$ the s-decay norm of the matrix in (2.17).

Lemma 2.4. *Let A be a Töplitz matrix as in (2.15), and $\mathfrak{s}_0 := (\nu + 2)/2$ (as defined above). Then*

$$|A(\varphi)|_s \leq C(\mathfrak{s}_0)|A|_{s + \mathfrak{s}_0}, \quad \forall \varphi \in \mathbb{T}^\nu.$$

PROOF. For all $\varphi \in \mathbb{T}^\nu$ we have

$$
\begin{aligned}
|A(\varphi)|^2_s &:= \sum_{j \in \mathbb{Z}} \langle j \rangle^{2s} \sup_{j_1 - j_2 = j} |A^{j_2}_{j_1}(\varphi)|^2 \prec \sum_{j \in \mathbb{Z}} \langle j \rangle^{2s} \sup_{j_1 - j_2 = j} \sum_{l \in \mathbb{Z}^\nu} |A^{j_2}_{j_1}(l)|^2 \langle l \rangle^{2\mathfrak{s}_0} \\
&\prec \sum_{j \in \mathbb{Z}} \sup_{j_1 - j_2 = j} \sum_{l \in \mathbb{Z}^\nu} |A^{j_2}_{j_1}(l)|^2 \langle l, j \rangle^{2(s + \mathfrak{s}_0)} \prec \sum_{j \in \mathbb{Z}, l \in \mathbb{Z}^\nu} \sup_{j_1 - j_2 = j} |A^{j_2}_{j_1}(l)|^2 \langle l, j \rangle^{2(s + \mathfrak{s}_0)} \\
&\overset{(2.16)}{\prec} |A|^2_{s + \mathfrak{s}_0},
\end{aligned}
$$

whence the lemma follows. ∎

Given $N \in \mathbb{N}$, we define the smoothing operator Π_N as

$$\left(\Pi_N A \right)^{(l_2, j_2)}_{(l_1, j_1)} := \begin{cases} A^{(l_2, j_2)}_{(l_1, j_1)} & \text{if } |l_1 - l_2| \leq N \\ 0 & \text{otherwise.} \end{cases} \tag{2.19}$$

Lemma 2.5. *The operator $\Pi_N^\perp := I - \Pi_N$ satisfies*

$$|\Pi_N^\perp A|_s \leq N^{-\beta}|A|_{s+\beta}, \quad |\Pi_N^\perp A|^{\mathrm{Lip}(\gamma)}_s \leq N^{-\beta}|A|^{\mathrm{Lip}(\gamma)}_{s+\beta}, \quad \beta \geq 0, \tag{2.20}$$

where in the second inequality $A := A(\lambda)$ is a Lipschitz family $\lambda \in \Lambda$.

2.2 Dynamical reducibility

All the transformations that we construct in sections 3 and 4 act on functions $u(\varphi, x)$ (of time and space). They can also be seen as:

(a) transformations of the phase space H^s_x that depend quasi-periodically on time (sections 3.1, 3.3-3.5 and 4);

(*b*) quasi-periodic reparametrizations of time (section 3.2).

This observation allows to interpret the conjugacy procedure from a dynamical point of view.
    Consider a quasi-periodic linear dynamical system

$$\partial_t u = L(\omega t)u. \tag{2.21}$$

We want to describe how (2.21) changes under the action of a transformation of type (*a*) or (*b*).
    Let $A(\omega t)$ be of type (*a*), and let $u = A(\omega t)v$. Then (2.21) is transformed into the linear system

$$\partial_t v = L_+(\omega t)v \quad \text{where} \quad L_+(\omega t) = A(\omega t)^{-1}L(\omega t)A(\omega t) - A(\omega t)^{-1}\partial_t A(\omega t). \tag{2.22}$$

The transformation $A(\omega t)$ may be regarded to act on functions $u(\varphi, x)$ as

$$(\tilde{A}u)(\varphi, x) := \big(A(\varphi)u(\varphi, \cdot)\big)(x) := A(\varphi)u(\varphi, x) \tag{2.23}$$

and one can check that $(\tilde{A}^{-1}u)(\varphi, x) = A^{-1}(\varphi)u(\varphi, x)$. The operator associated to (2.21) (on quasi-periodic functions)

$$\mathcal{L} := \omega \cdot \partial_\varphi - L(\varphi) \tag{2.24}$$

transforms under the action of $\tilde{A}$ into

$$\tilde{A}^{-1}\mathcal{L}\tilde{A} = \omega \cdot \partial_\varphi - L_+(\varphi),$$

which is exactly the linear system in (2.22), acting on quasi-periodic functions.
    Now consider a transformation of type (*b*), namely a change of the time variable

$$\tau := t + \alpha(\omega t) \iff t = \tau + \tilde{\alpha}(\omega\tau); \quad (Bv)(t) := v(t + \alpha(\omega t)), \quad (B^{-1}u)(\tau) = u(\tau + \tilde{\alpha}(\omega\tau)), \tag{2.25}$$

where $\alpha = \alpha(\varphi)$, $\varphi \in \mathbb{T}^\nu$, is a 2π-periodic function of ν variables (in other words, $t \mapsto t + \alpha(\omega t)$ is the diffeomorphisms of $\mathbb{R}$ induced by the transformation B). If $u(t)$ is a solution of (2.21), then $v(\tau)$, defined by $u = Bv$, solves

$$\partial_\tau v(\tau) = L_+(\omega\tau)v(\tau), \quad L_+(\omega\tau) := \left(\frac{L(\omega t)}{1 + (\omega \cdot \partial_\varphi\alpha)(\omega t)}\right)_{|t=\tau+\tilde{\alpha}(\omega\tau)}. \tag{2.26}$$

We may regard the associated transformation on quasi-periodic functions defined by

$$(\tilde{B}h)(\varphi, x) := h(\varphi + \omega\alpha(\varphi), x), \quad (\tilde{B}^{-1}h)(\varphi, x) := h(\varphi + \omega\tilde{\alpha}(\varphi), x),$$

as in step 3.2, where we calculate

$$B^{-1}\mathcal{L}B = \rho(\varphi)\mathcal{L}_+, \quad \rho(\varphi) := B^{-1}(1 + \omega \cdot \partial_\varphi\alpha),$$

$$\mathcal{L}_+ = \omega \cdot \partial_\varphi - L_+(\varphi), \quad L_+(\varphi) := \frac{1}{\rho(\varphi)}L(\varphi + \omega\tilde{\alpha}(\varphi)). \tag{2.27}$$

(2.27) is nothing but the linear system (2.26), acting on quasi-periodic functions.

2.3   Real, reversible and Hamiltonian operators

We consider the space of *real* functions

$$Z := \{u(\varphi, x) = \overline{u(\varphi, x)}\}, \tag{2.28}$$

and of *even* (in space-time), respectively *odd*, functions

$$X := \{u(\varphi, x) = u(-\varphi, -x)\}, \quad Y := \{u(\varphi, x) = -u(-\varphi, -x)\}. \tag{2.29}$$

Definition 2.2. *An operator R is*

1. REAL *if* $R : Z \to Z$

2. REVERSIBLE *if* $R : X \to Y$

3. REVERSIBILITY-PRESERVING *if* $R : X \to X$, $R : Y \to Y$.

The composition of a reversible and a reversibility-preserving operator is reversible.
The above properties may be characterized in terms of matrix elements.

Lemma 2.6. *We have*

$$R : X \to Y \iff R^{-j}_{-k}(-l) = -R^j_k(l), \qquad R : X \to X \iff R^{-j}_{-k}(-l) = R^j_k(l),$$

$$R : Z \to Z \iff \overline{R^j_k(l)} = R^{-j}_{-k}(-l).$$

For the Hamiltonian KdV the phase space is $H^1_0 := \{u \in H^1(\mathbb{T}) \ : \ \int_{\mathbb{T}} u(x)dx = 0\}$ and it is more convenient the dynamical systems perspective.

Definition 2.3. *A time dependent linear vector field* $X(t) : H^1_0 \to H^1_0$ *is* HAMILTONIAN *if* $X(t) = \partial_x G(t)$ *for some real linear operator $G(t)$ which is self-adjoint with respect to the L^2 scalar product.*

If $G(t) = G(\omega t)$ is quasi-periodic in time, we say that the associated operator $\omega \cdot \partial_\varphi - \partial_x G(\varphi)$ (see (2.24))
is Hamiltonian.

Definition 2.4. *A map* $A : H^1_0 \to H^1_0$ *is* SYMPLECTIC *if*

$$\Omega(Au, Av) = \Omega(u, v), \quad \forall u, v \in H^1_0, \tag{2.30}$$

where the symplectic 2-form Ω is defined in (1.12). Equivalently $A^T \partial^{-1}_x A = \partial^{-1}_x$.

If $A(\varphi)$, $\forall \varphi \in \mathbb{T}^\nu$, is a family of symplectic maps we say that the corresponding operator in (2.23) is
symplectic.

Under a time dependent family of symplectic transformations $u = \Phi(t)v$ the linear Hamiltonian equation

$$u_t = \partial_x G(t)u \quad \text{with Hamiltonian} \quad H(t, u) := \tfrac{1}{2} \big(G(t)u, u\big)_{L^2}$$

transforms into the equation

$$v_t = \partial_x E(t)v, \quad E(t) := \Phi(t)^T G(t)\Phi(t) - \Phi(t)^T \partial^{-1}_x \Phi_t(t)$$

with Hamiltonian

$$K(t, v) = \tfrac{1}{2} \big(G(t)\Phi(t)v, \Phi(t)v\big)_{L^2} - \tfrac{1}{2} \big(\partial^{-1}_x \Phi_t(t)v, \Phi(t)v\big)_{L^2}. \tag{2.31}$$

Note that $E(t)$ is self-adjoint with respect to the L^2 scalar product because $\Phi^T \partial^{-1}_x \Phi_t + \Phi^T_t \partial^{-1}_x \Phi = 0$.

3 Regularization of the linearized operator

Our existence proof is based on a Nash-Moser iterative scheme. The main step concerns the invertibility of the linearized operator (see (1.16))

$$\mathcal{L}h = \mathcal{L}(\lambda, u, \varepsilon)h := \omega \cdot \partial_\varphi h + (1 + a_3)\partial_{xxx}h + a_2\partial_{xx}h + a_1\partial_x h + a_0 h \tag{3.1}$$

obtained linearizing (1.4) at any approximate (or exact) solution u. The coefficients $a_i = a_i(\varphi, x) = a_i(u, \varepsilon)(\varphi, x)$ are periodic functions of (φ, x), depending on u, ε. They are explicitly obtained from the partial derivatives of $\varepsilon f(\varphi, x, z)$ as

$$a_i(\varphi, x) = \varepsilon(\partial_{z_i} f)\big(\varphi, x, u(\varphi, x), u_x(\varphi, x), u_{xx}(\varphi, x), u_{xxx}(\varphi, x)\big), \quad i = 0, 1, 2, 3. \tag{3.2}$$

The operator $\mathcal{L}$ depends on λ because $\omega = \lambda\bar\omega$. Since ε is a (small) fixed parameter, we simply write $\mathcal{L}(\lambda, u)$ instead of $\mathcal{L}(\lambda, u, \varepsilon)$, and $a_i(u)$ instead of $a_i(u, \varepsilon)$. We emphasize that the coefficients a_i do not depend explicitly on the parameter λ (they depend on λ only through $u(\lambda)$).

In the Hamiltonian case (1.11) the linearized KdV operator (3.1) has the form

$$\mathcal{L}h = \omega \cdot \partial_\varphi h + \partial_x\Big(\partial_x\big\{A_1(\varphi, x)\partial_x h\big\} - A_0(\varphi, x)h\Big)$$

where

$$A_1(\varphi, x) := 1 + \varepsilon(\partial_{z_1 z_1} F)(\varphi, x, u, u_x), \quad A_0(\varphi, x) := -\varepsilon\partial_x\{(\partial_{z_0 z_1} F)(\varphi, x, u, u_x)\} + \varepsilon(\partial_{z_0 z_0} F)(\varphi, x, u, u_x)$$

and it is generated by the quadratic Hamiltonian

$$H_L(\varphi, h) := \frac{1}{2}\int_{\mathbb{T}}\Big(A_0(\varphi, x)h^2 + A_1(\varphi, x)h_x^2\Big)\,dx, \quad h \in H_0^1.$$

Remark 3.1. *In the reversible case, i.e. the nonlinearity f satisfies (1.13) and $u \in X$ (see (2.29), (1.14)) the coefficients a_i satisfy the parity*

$$a_3, a_1 \in X, \quad a_2, a_0 \in Y, \tag{3.3}$$

and $\mathcal{L}$ maps X into Y, namely $\mathcal{L}$ is reversible, see Definition 2.2.

Remark 3.2. *In the Hamiltonian case (1.11), assumption (Q)-(1.7) is automatically satisfied (with $\alpha(\varphi) = 2$) because*

$$f(\varphi, x, u, u_x, u_{xx}, u_{xxx}) = a(\varphi, x, u, u_x) + b(\varphi, x, u, u_x)u_{xx} + c(\varphi, x, u, u_x)u_{xx}^2 + d(\varphi, x, u, u_x)u_{xxx}$$

where

$$b = 2(\partial_{z_1 z_1 x}^3 F) + 2z_1(\partial_{z_1 z_1 z_0}^3 F), \qquad c = \partial_{z_1}^3 F, \qquad d = \partial_{z_1}^2 F,$$

and so

$$\partial_{z_2} f = b + 2z_2 c = 2(d_x + z_1 d_{z_0} + z_2 d_{z_1}) = 2\Big(\partial_{z_3 x}^2 f + z_1 \partial_{z_3 z_0}^2 f + z_2 \partial_{z_3 z_1}^2 f + z_3 \partial_{z_3 z_2}^2 f\Big).$$

The coefficients a_i, together with their derivative $\partial_u a_i(u)[h]$ with respect to u in the direction h, satisfy tame estimates:

Lemma 3.1. *Let $f \in C^q$, see (1.3). For all $\mathfrak{s}_0 \leq s \leq q - 2$, $\|u\|_{\mathfrak{s}_0+3} \leq 1$, we have, for all $i = 0, 1, 2, 3$,*

$$\|a_i(u)\|_s \leq \varepsilon\, C(s)\big(1 + \|u\|_{s+3}\big), \tag{3.4}$$

$$\|\partial_u a_i(u)[h]\|_s \leq \varepsilon\, C(s)\big(\|h\|_{s+3} + \|u\|_{s+3}\|h\|_{\mathfrak{s}_0+3}\big). \tag{3.5}$$

If, moreover, $\lambda \mapsto u(\lambda) \in H^s$ is Lipschitz family satisfying $\|u\|_{\mathfrak{s}_0+3}^{\mathrm{Lip}(\gamma)} \leq 1$ (see (2.2)), then

$$\|a_i\|_s^{\mathrm{Lip}(\gamma)} \leq \varepsilon\, C(s)\big(1 + \|u\|_{s+3}^{\mathrm{Lip}(\gamma)}\big). \tag{3.6}$$

PROOF. The tame estimate (3.4) follows by Lemma 6.2(i) applied to the function $\partial_{z_i} f$, $i = 0, \ldots, 3$, which is valid for $s + 1 \leq q$. The tame bound (3.5) for

$$\partial_u a_i(u)[h] \overset{(3.2)}{=} \varepsilon\sum_{k=0}^{3}(\partial_{z_k z_i}^2 f)\big(\varphi, x, u, u_x, u_{xx}, u_{xxx}\big)\partial_x^k h, \quad i = 0, \ldots, 3,$$

follows by (6.5) and applying Lemma 6.2(i) to the functions $\partial_{z_k z_i}^2 f$, which gives

$$\|(\partial_{z_k z_i}^2 f)\big(\varphi, x, u, u_x, u_{xx}, u_{xxx}\big)\|_s \leq C(s)\|f\|_{C^{s+2}}(1 + \|u\|_{s+3}),$$

for $s + 2 \leq q$. The Lipschitz bound (3.6) follows similarly. ∎

3.1 Step 1. Change of the space variable

We consider a φ-dependent family of diffeomorphisms of the 1-dimensional torus $\mathbb{T}$ of the form

$$y = x + \beta(\varphi, x), \tag{3.7}$$

where β is a (small) real-valued function, 2π periodic in all its arguments. The change of variables (3.7) induces on the space of functions the linear operator

$$(\mathcal{A}h)(\varphi, x) := h(\varphi, x + \beta(\varphi, x)). \tag{3.8}$$

The operator $\mathcal{A}$ is invertible, with inverse

$$(\mathcal{A}^{-1}v)(\varphi, y) = v(\varphi, y + \tilde{\beta}(\varphi, y)), \tag{3.9}$$

where $y \mapsto y + \tilde{\beta}(\varphi, y)$ is the inverse diffeomorphism of (3.7), namely

$$x = y + \tilde{\beta}(\varphi, y) \quad \Longleftrightarrow \quad y = x + \beta(\varphi, x). \tag{3.10}$$

Remark 3.3. *In the Hamiltonian case (1.11) we use, instead of (3.8), the modified change of variable (1.25) which is symplectic, for each $\varphi \in \mathbb{T}^\nu$. Indeed, setting $U := \partial_x^{-1} u$ (and neglecting to write the φ-dependence)*

$$\Omega(\mathcal{A}u, \mathcal{A}v) = \int_{\mathbb{T}} \partial_x^{-1}\Big(\partial_x\big\{U(x + \beta(x))\big\}\Big)(1 + \beta_x(x))v(x + \beta(x))\,dx$$

$$= \int_{\mathbb{T}} U(x + \beta(x))(1 + \beta_x(x))v(x + \beta(x))dx - c\int_{\mathbb{T}}(1 + \beta_x(x))v(x + \beta(x))dx$$

$$= \int_{\mathbb{T}} U(y)v(y)dy = \Omega(u, v), \quad v \in H_0^1,$$

where c is the average of $U(x + \beta(x))$ in $\mathbb{T}$. The inverse operator of (1.25) is $(\mathcal{A}^{-1}v)(\varphi, y) = (1 + \tilde{\beta}_y(\varphi, y))v(y + \tilde{\beta}(\varphi, y))$ which is also symplectic.

Now we calculate the conjugate $\mathcal{A}^{-1}\mathcal{L}\mathcal{A}$ of the linearized operator $\mathcal{L}$ in (3.1) with $\mathcal{A}$ in (3.8).

The conjugate $\mathcal{A}^{-1}a\mathcal{A}$ of any multiplication operator $a : h(\varphi, x) \mapsto a(\varphi, x)h(\varphi, x)$ is the multiplication operator $(\mathcal{A}^{-1}a)$ that maps $v(\varphi, y) \mapsto (\mathcal{A}^{-1}a)(\varphi, y)\,v(\varphi, y)$. By conjugation, the differential operators become

$$\mathcal{A}^{-1}\omega \cdot \partial_\varphi \mathcal{A} = \omega \cdot \partial_\varphi + \{\mathcal{A}^{-1}(\omega \cdot \partial_\varphi \beta)\}\,\partial_y,$$
$$\mathcal{A}^{-1}\partial_x \mathcal{A} = \{\mathcal{A}^{-1}(1 + \beta_x)\}\,\partial_y,$$
$$\mathcal{A}^{-1}\partial_{xx}\mathcal{A} = \{\mathcal{A}^{-1}(1 + \beta_x)^2\}\,\partial_{yy} + \{\mathcal{A}^{-1}(\beta_{xx})\}\,\partial_y,$$
$$\mathcal{A}^{-1}\partial_{xxx}\mathcal{A} = \{\mathcal{A}^{-1}(1 + \beta_x)^3\}\,\partial_{yyy} + \{3\mathcal{A}^{-1}[(1 + \beta_x)\beta_{xx}]\}\,\partial_{yy} + \{\mathcal{A}^{-1}(\beta_{xxx})\}\,\partial_y,$$

where all the coefficients $\{\mathcal{A}^{-1}(\ldots)\}$ are periodic functions of (φ, y). Thus (recall (3.1))

$$\mathcal{L}_1 := \mathcal{A}^{-1}\mathcal{L}\mathcal{A} = \omega \cdot \partial_\varphi + b_3(\varphi, y)\partial_{yyy} + b_2(\varphi, y)\partial_{yy} + b_1(\varphi, y)\partial_y + b_0(\varphi, y) \tag{3.11}$$

where

$$b_3 = \mathcal{A}^{-1}[(1 + a_3)(1 + \beta_x)^3], \qquad b_1 = \mathcal{A}^{-1}[\omega \cdot \partial_\varphi \beta + (1 + a_3)\beta_{xxx} + a_2\beta_{xx} + a_1(1 + \beta_x)], \tag{3.12}$$

$$b_0 = \mathcal{A}^{-1}(a_0), \qquad b_2 = \mathcal{A}^{-1}[(1 + a_3)3(1 + \beta_x)\beta_{xx} + a_2(1 + \beta_x)^2]. \tag{3.13}$$

We look for $\beta(\varphi, x)$ such that the coefficient $b_3(\varphi, y)$ of the highest order derivative ∂_{yyy} in (3.11) does not depend on y, namely

$$b_3(\varphi, y) \overset{(3.12)}{=} \mathcal{A}^{-1}[(1 + a_3)(1 + \beta_x)^3](\varphi, y) = b(\varphi) \tag{3.14}$$

for some function $b(\varphi)$ of φ only. Since $\mathcal{A}$ changes only the space variable, $\mathcal{A}b = b$ for every function $b(\varphi)$ that is independent on y. Hence (3.14) is equivalent to

$$\big(1 + a_3(\varphi, x)\big)\big(1 + \beta_x(\varphi, x)\big)^3 = b(\varphi), \tag{3.15}$$

15

namely

$$\beta_x = \rho_0, \qquad \rho_0(\varphi, x) := b(\varphi)^{1/3}\big(1 + a_3(\varphi, x)\big)^{-1/3} - 1. \tag{3.16}$$

The equation (3.16) has a solution β, periodic in x, if and only if $\int_{\mathbb{T}} \rho_0(\varphi, x)\, dx = 0$. This condition uniquely determines

$$b(\varphi) = \left(\frac{1}{2\pi} \int_{\mathbb{T}} \big(1 + a_3(\varphi, x)\big)^{-\frac{1}{3}}\, dx \right)^{-3}. \tag{3.17}$$

Then we fix the solution (with zero average) of (3.16),

$$\beta(\varphi, x) := (\partial_x^{-1} \rho_0)(\varphi, x), \tag{3.18}$$

where ∂_x^{-1} is defined by linearity as

$$\partial_x^{-1} e^{\mathrm{i}jx} := \frac{e^{\mathrm{i}jx}}{\mathrm{i}j} \quad \forall j \in \mathbb{Z} \setminus \{0\}, \qquad \partial_x^{-1} 1 = 0. \tag{3.19}$$

In other words, $\partial_x^{-1} h$ is the primitive of h with zero average in x.

With this choice of β, we get (see (3.11), (3.14))

$$\mathcal{L}_1 = \mathcal{A}^{-1}\mathcal{L}\mathcal{A} = \omega \cdot \partial_\varphi + b_3(\varphi)\partial_{yyy} + b_2(\varphi, y)\partial_{yy} + b_1(\varphi, y)\partial_y + b_0(\varphi, y), \tag{3.20}$$

where $b_3(\varphi) := b(\varphi)$ is defined in (3.17).

Remark 3.4. *In the reversible case, $\beta \in Y$ because $a_3 \in X$, see (3.3). Therefore the operator A in (3.8), as well as $\mathcal{A}^{-1}$ in (3.9), maps $X \to X$ and $Y \to Y$, namely it is reversibility-preserving, see Definition 2.2. By (3.3) the coefficients of $\mathcal{L}_1$ (see (3.12), (3.13)) have parity*

$$b_3, b_1 \in X, \qquad b_2, b_0 \in Y, \tag{3.21}$$

and $\mathcal{L}_1$ maps $X \to Y$, namely it is reversible.

Remark 3.5. *In the Hamiltonian case (1.11) the resulting operator $\mathcal{L}_1$ in (3.20) is Hamiltonian and $b_2(\varphi, y) = 2\partial_y b_3(\varphi) \equiv 0$. Actually, by (2.31), the corresponding Hamiltonian has the form*

$$K(\varphi, v) = \frac{1}{2} \int_{\mathbb{T}} b_3(\varphi) v_y^2 + B_0(\varphi, y) v^2\, dy, \tag{3.22}$$

for some function $B_0(\varphi, y)$.

3.2 Step 2. Time reparametrization

The goal of this section is to make constant the coefficient of the highest order spatial derivative operator ∂_{yyy} of $\mathcal{L}_1$ in (3.20), by a quasi-periodic reparametrization of time. We consider a diffeomorphism of the torus $\mathbb{T}^\nu$ of the form

$$\varphi \mapsto \varphi + \omega\alpha(\varphi), \quad \varphi \in \mathbb{T}^\nu, \quad \alpha(\varphi) \in \mathbb{R}, \tag{3.23}$$

where α is a (small) *real* valued function, 2π-periodic in all its arguments. The induced linear operator on the space of functions is

$$(Bh)(\varphi, y) := h\big(\varphi + \omega\alpha(\varphi),\, y\big) \tag{3.24}$$

whose inverse is

$$(B^{-1}v)(\vartheta, y) := v\big(\vartheta + \omega\tilde{\alpha}(\vartheta),\, y\big) \tag{3.25}$$

where $\varphi = \vartheta + \omega\tilde{\alpha}(\vartheta)$ is the inverse diffeomorphism of $\vartheta = \varphi + \omega\alpha(\varphi)$. By conjugation, the differential operators become

$$B^{-1}\omega \cdot \partial_\vartheta B = \rho(\vartheta)\,\omega \cdot \partial_\vartheta, \quad B^{-1}\partial_y B = \partial_y, \quad \rho := B^{-1}(1 + \omega \cdot \partial_\varphi \alpha). \tag{3.26}$$

Thus, see (3.20),

$$B^{-1}\mathcal{L}_1 B = \rho\,\omega\cdot\partial_\vartheta + \{B^{-1}b_3\}\,\partial_{yyy} + \{B^{-1}b_2\}\,\partial_{yy} + \{B^{-1}b_1\}\,\partial_y + \{B^{-1}b_0\}. \tag{3.27}$$

We look for $\alpha(\varphi)$ such that the (variable) coefficients of the highest order derivatives ($\omega\cdot\partial_\vartheta$ and ∂_{yyy}) are proportional, namely

$$\{B^{-1}b_3\}(\vartheta) = m_3\rho(\vartheta) = m_3\{B^{-1}(1+\omega\cdot\partial_\varphi\alpha)\}(\vartheta) \tag{3.28}$$

for some constant $m_3 \in \mathbb{R}$. Since B is invertible, this is equivalent to require that

$$b_3(\varphi) = m_3\big(1 + \omega\cdot\partial_\varphi\alpha(\varphi)\big). \tag{3.29}$$

Integrating on $\mathbb{T}^\nu$ determines the value of the constant m_3,

$$m_3 := \frac{1}{(2\pi)^\nu}\int_{\mathbb{T}^\nu} b_3(\varphi)\,d\varphi. \tag{3.30}$$

Thus we choose the unique solution of (3.29) with zero average

$$\alpha(\varphi) := \frac{1}{m_3}\,(\omega\cdot\partial_\varphi)^{-1}(b_3 - m_3)(\varphi) \tag{3.31}$$

where $(\omega\cdot\partial_\varphi)^{-1}$ is defined by linearity

$$(\omega\cdot\partial_\varphi)^{-1}e^{\mathrm{i}l\cdot\varphi} := \frac{e^{\mathrm{i}l\cdot\varphi}}{\mathrm{i}\omega\cdot l}\,,\ l \neq 0\,, \quad (\omega\cdot\partial_\varphi)^{-1}1 = 0\,.$$

With this choice of α we get (see (3.27), (3.28))

$$B^{-1}\mathcal{L}_1 B = \rho\,\mathcal{L}_2, \qquad \mathcal{L}_2 := \omega\cdot\partial_\vartheta + m_3\,\partial_{yyy} + c_2(\vartheta,y)\,\partial_{yy} + c_1(\vartheta,y)\,\partial_y + c_0(\vartheta,y), \tag{3.32}$$

where

$$c_i := \frac{B^{-1}b_i}{\rho}, \quad i = 0,1,2. \tag{3.33}$$

Remark 3.6. *In the reversible case, α is odd because b_3 is even (see (3.21)), and B is reversibility preserving. Since ρ (defined in (3.26)) is even, the coefficients $c_3, c_1 \in X$, $c_2, c_0 \in Y$ and $\mathcal{L}_2 : X \to Y$ is reversible.*

Remark 3.7. *In the Hamiltonian case, the operator $\mathcal{L}_2$ is still Hamiltonian (the new Hamiltonian is the old one at the new time, divided by the factor ρ). The coefficient $c_2(\vartheta,y) \equiv 0$ because $b_2 \equiv 0$, see remark 3.5.*

3.3 Step 3. Descent method: step zero

The aim of this section is to eliminate the term of order ∂_{yy} from $\mathcal{L}_2$ in (3.32).

Consider the multiplication operator

$$\mathcal{M}h := v(\vartheta,y)h \tag{3.34}$$

where the function v is periodic in all its arguments. Calculate the difference

$$\mathcal{L}_2\,\mathcal{M} - \mathcal{M}\,(\omega\cdot\partial_\vartheta + m_3\partial_{yyy}) = T_2\partial_{yy} + T_1\partial_y + T_0, \tag{3.35}$$

where

$$T_2 := 3m_3v_y + c_2v, \quad T_1 := 3m_3v_{yy} + 2c_2v_y + c_1v, \quad T_0 := \omega\cdot\partial_\vartheta v + m_3v_{yyy} + c_2v_{yy} + c_1v_y + c_0v. \tag{3.36}$$

To eliminate the factor T_2, we need

$$3m_3v_y + c_2v = 0. \tag{3.37}$$

Equation (3.37) has the periodic solution

$$v(\vartheta,y) = \exp\left\{-\frac{1}{3m_3}\,(\partial_y^{-1}c_2)(\vartheta,y)\right\} \tag{3.38}$$

provided that

$$\int_{\mathbb{T}} c_2(\vartheta, y)\, dy = 0. \tag{3.39}$$

Let us prove (3.39). By (3.33), (3.26), for each $\vartheta = \varphi + \omega\alpha(\varphi)$ we get

$$\int_{\mathbb{T}} c_2(\vartheta, y)\, dy = \frac{1}{\{B^{-1}(1 + \omega \cdot \partial_\varphi \alpha)\}(\vartheta)} \int_{\mathbb{T}} (B^{-1} b_2)(\vartheta, y)\, dy = \frac{1}{1 + \omega \cdot \partial_\varphi \alpha(\varphi)} \int_{\mathbb{T}} b_2(\varphi, y)\, dy.$$

By the definition (3.13) of b_2 and changing variable $y = x + \beta(\varphi, x)$ in the integral (recall (3.8))

$$\int_{\mathbb{T}} b_2(\varphi, y)\, dy \overset{(3.13)}{=} \int_{\mathbb{T}} \left((1 + a_3) 3 (1 + \beta_x)\beta_{xx} + a_2(1 + \beta_x)^2 \right)(1 + \beta_x)\, dx$$

$$\overset{(3.15)}{=} b(\varphi) \left\{ 3 \int_{\mathbb{T}} \frac{\beta_{xx}(\varphi, x)}{1 + \beta_x(\varphi, x)}\, dx + \int_{\mathbb{T}} \frac{a_2(\varphi, x)}{1 + a_3(\varphi, x)}\, dx \right\}. \tag{3.40}$$

The first integral in (3.40) is zero because $\beta_{xx}/(1 + \beta_x) = \partial_x \log(1 + \beta_x)$. The second one is zero because of assumptions (Q)-(1.7) or (F)-(1.6), see (1.26). As a consequence (3.39) is proved, and (3.37) has the periodic solution v defined in (3.38). Note that v is close to 1 for ε small. Hence the multiplication operator $\mathcal{M}$ defined in (3.34) is invertible and $\mathcal{M}^{-1}$ is the multiplication operator for $1/v$. By (3.35) and since $T_2 = 0$, we deduce

$$\mathcal{L}_3 := \mathcal{M}^{-1} \mathcal{L}_2 \mathcal{M} = \omega \cdot \partial_\vartheta + m_3 \partial_{yyy} + d_1(\vartheta, y)\partial_y + d_0(\vartheta, y), \qquad d_i := \frac{T_i}{v}, \quad i = 0, 1. \tag{3.41}$$

Remark 3.8. *In the reversible case, since c_2 is odd (see Remark 3.6) the function v is even, then $\mathcal{M}$, $\mathcal{M}^{-1}$ are reversibility preserving and by (3.36) and (3.41) $d_1 \in X$ and $d_0 \in Y$, which implies that $\mathcal{L}_3 : X \to Y$.*

Remark 3.9. *In the Hamiltonian case, there is no need to perform this step because $c_2 \equiv 0$, see remark 3.7.*

3.4 Step 4. Change of space variable (translation)

Consider the change of the space variable

$$z = y + p(\vartheta)$$

which induces the operators

$$\mathcal{T} h(\vartheta, y) := h(\vartheta, y + p(\vartheta)), \quad \mathcal{T}^{-1} v(\vartheta, z) := v(\vartheta, z - p(\vartheta)). \tag{3.42}$$

The differential operators become

$$\mathcal{T}^{-1} \omega \cdot \partial_\vartheta \mathcal{T} = \omega \cdot \partial_\vartheta + \{\omega \cdot \partial_\vartheta p(\vartheta)\} \partial_z, \qquad \mathcal{T}^{-1} \partial_y \mathcal{T} = \partial_z.$$

Thus, by (3.41),

$$\mathcal{L}_4 := \mathcal{T}^{-1} \mathcal{L}_3 \mathcal{T} = \omega \cdot \partial_\vartheta + m_3 \partial_{zzz} + e_1(\vartheta, z) \partial_z + e_0(\vartheta, z)$$

where

$$e_1(\vartheta, z) := \omega \cdot \partial_\vartheta p(\vartheta) + (\mathcal{T}^{-1} d_1)(\vartheta, z), \quad e_0(\vartheta, z) := (\mathcal{T}^{-1} d_0)(\vartheta, z). \tag{3.43}$$

Now we look for $p(\vartheta)$ such that the average

$$\frac{1}{2\pi} \int_{\mathbb{T}} e_1(\vartheta, z)\, dz = m_1, \quad \forall \vartheta \in \mathbb{T}^\nu, \tag{3.44}$$

for some constant $m_1 \in \mathbb{R}$ (independent of ϑ). Equation (3.44) is equivalent to

$$\omega \cdot \partial_\vartheta p = m_1 - \int_{\mathbb{T}} d_1(\vartheta, y)\, dy =: V(\vartheta). \tag{3.45}$$

18

The equation (3.45) has a periodic solution $p(\vartheta)$ if and only if $\int_{\mathbb{T}^\nu} V(\vartheta)\,d\vartheta = 0$. Hence we have to define

$$m_1 := \frac{1}{(2\pi)^{\nu+1}} \int_{\mathbb{T}^{\nu+1}} d_1(\vartheta, y)\,d\vartheta dy \tag{3.46}$$

and

$$p(\vartheta) := (\omega \cdot \partial_\vartheta)^{-1} V(\vartheta). \tag{3.47}$$

With this choice of p, after renaming the space-time variables $z = x$ and $\vartheta = \varphi$, we have

$$\mathcal{L}_4 = \omega \cdot \partial_\varphi + m_3 \partial_{xxx} + e_1(\varphi, x)\,\partial_x + e_0(\varphi, x), \qquad \frac{1}{2\pi} \int_{\mathbb{T}} e_1(\varphi, x)\,dx = m_1, \quad \forall \varphi \in \mathbb{T}^\nu. \tag{3.48}$$

Remark 3.10. *By (3.45), (3.47) and since $d_1 \in X$ (see remark 3.8), the function p is odd. Then $\mathcal{T}$ and $\mathcal{T}^{-1}$ defined in (3.42) are reversibility preserving and the coefficients e_1, e_0 defined in (3.43) satisfy $e_1 \in X$, $e_0 \in Y$. Hence $\mathcal{L}_4 : X \to Y$ is reversible.*

Remark 3.11. *In the Hamiltonian case the operator $\mathcal{L}_4$ is Hamiltonian, because the operator $\mathcal{T}$ in (3.42) is symplectic (it is a particular case of the change of variables (1.25) with $\beta(\varphi, x) = p(\varphi)$).*

3.5 Step 5. Descent method: conjugation by pseudo-differential operators

The goal of this section is to conjugate $\mathcal{L}_4$ in (3.48) to an operator of the form $\omega \cdot \partial_\varphi + m_3 \partial_{xxx} + m_1 \partial_x + \mathcal{R}$ where the constants m_3, m_1 are defined in (3.30), (3.46), and $\mathcal{R}$ is a pseudo-differential operator of order 0.

Consider an operator of the form

$$\mathcal{S} := I + w(\varphi, x) \partial_x^{-1} \tag{3.49}$$

where $w : \mathbb{T}^{\nu+1} \to \mathbb{R}$ and the operator ∂_x^{-1} is defined in (3.19). Note that $\partial_x^{-1}\partial_x = \partial_x\partial_x^{-1} = \pi_0$, where π_0 is the L^2-projector on the subspace $H_0 := \{u(\varphi, x) \in L^2(\mathbb{T}^{\nu+1}) : \int_{\mathbb{T}} u(\varphi, x)\,dx = 0\}$.

A direct computation shows that the difference

$$\mathcal{L}_4 \mathcal{S} - \mathcal{S}(\omega \cdot \partial_\varphi + m_3 \partial_{xxx} + m_1 \partial_x) = r_1 \partial_x + r_0 + r_{-1} \partial_x^{-1} \tag{3.50}$$

where (using $\partial_x \pi_0 = \pi_0 \partial_x = \partial_x$, $\partial_x^{-1} \partial_{xxx} = \partial_{xx}$)

$$r_1 := 3m_3 w_x + e_1(\varphi, x) - m_1 \tag{3.51}$$

$$r_0 := e_0 + \big(3m_3 w_{xx} + e_1 w - m_1 w\big)\pi_0 \tag{3.52}$$

$$r_{-1} := \omega \cdot \partial_\varphi w + m_3 w_{xxx} + e_1 w_x. \tag{3.53}$$

We look for a periodic function $w(\varphi, x)$ such that $r_1 = 0$. By (3.51) and (3.44) we take

$$w = \frac{1}{3m_3} \partial_x^{-1}[m_1 - e_1]. \tag{3.54}$$

For ε small enough the operator $\mathcal{S}$ is invertible and we obtain, by (3.50),

$$\mathcal{L}_5 := \mathcal{S}^{-1} \mathcal{L}_4 \mathcal{S} = \omega \cdot \partial_\varphi + m_3 \partial_{xxx} + m_1 \partial_x + \mathcal{R}, \qquad \mathcal{R} := \mathcal{S}^{-1}(r_0 + r_{-1}\partial_x^{-1}). \tag{3.55}$$

Remark 3.12. *In the reversible case, the function $w \in Y$, because $e_1 \in X$, see remark 3.10. Then $\mathcal{S}, \mathcal{S}^{-1}$ are reversibility preserving. By (3.52) and (3.53), $r_0 \in Y$ and $r_{-1} \in X$. Then the operators $\mathcal{R}, \mathcal{L}_5$ defined in (3.55) are reversible, namely $\mathcal{R}, \mathcal{L}_5 : X \to Y$.*

Remark 3.13. *In the Hamiltonian case, we consider, instead of (3.49), the modified operator*

$$\mathcal{S} := e^{\pi_0 w(\varphi, x)\partial_x^{-1}} := I + \pi_0 w(\varphi, x)\partial_x^{-1} + \ldots \tag{3.56}$$

which, for each $\varphi \in \mathbb{T}^\nu$, is symplectic. Actually $\mathcal{S}$ is the time one flow map of the Hamiltonian vector field $\pi_0 w(\varphi, x)\partial_x^{-1}$ which is generated by the Hamiltonian

$$H_{\mathcal{S}}(\varphi, u) := -\frac{1}{2} \int_{\mathbb{T}} w(\varphi, x)\big(\partial_x^{-1} u\big)^2 dx, \quad u \in H_0^1.$$

The corresponding $\mathcal{L}_5$ in (3.55) is Hamiltonian. Note that the operators (3.56) and (3.49) differ only for pseudo-differential smoothing operators of order $O(\partial_x^{-2})$ and of smaller size $O(w^2) = O(\varepsilon^2)$.

3.6 Estimates on $\mathcal{L}_5$

Summarizing the steps performed in the previous sections 3.1-3.5, we have (semi)-conjugated the operator $\mathcal{L}$ defined in (3.1) to the operator $\mathcal{L}_5$ defined in (3.55), namely

$$\mathcal{L} = \Phi_1 \mathcal{L}_5 \Phi_2^{-1}, \qquad \Phi_1 := \mathcal{AB}\rho\mathcal{MTS}, \quad \Phi_2 := \mathcal{ABMTS} \tag{3.57}$$

(where ρ means the multiplication operator for the function ρ defined in (3.26)).

In the next lemma we give tame estimates for $\mathcal{L}_5$ and Φ_1, Φ_2. We define the constants

$$\sigma := 2\tau_0 + 2\nu + 17, \quad \sigma' := 2\tau_0 + \nu + 14 \tag{3.58}$$

where τ_0 is defined in (1.2) and ν is the number of frequencies.

Lemma 3.2. *Let $f \in C^q$, see (1.3), and $\mathfrak{s}_0 \le s \le q - \sigma$. There exists $\delta > 0$ such that, if $\varepsilon\gamma_0^{-1} < \delta$ (the constant γ_0 is defined in (1.2)), then, for all*

$$\|u\|_{\mathfrak{s}_0+\sigma} \le 1, \tag{3.59}$$

(i) the transformations Φ_1, Φ_2 defined in (3.57) are invertible operators of $H^s(\mathbb{T}^{\nu+1})$, and satisfy

$$\|\Phi_i h\|_s + \|\Phi_i^{-1} h\|_s \le C(s)\big(\|h\|_s + \|u\|_{s+\sigma}\|h\|_{\mathfrak{s}_0}\big), \tag{3.60}$$

for $i = 1, 2$. Moreover, if $u(\lambda), h(\lambda)$ are Lipschitz families with

$$\|u\|_{\mathfrak{s}_0+\sigma}^{\mathrm{Lip}(\gamma)} \le 1, \tag{3.61}$$

then

$$\|\Phi_i h\|_s^{\mathrm{Lip}(\gamma)} + \|\Phi_i^{-1} h\|_s^{\mathrm{Lip}(\gamma)} \le C(s)\big(\|h\|_{s+3}^{\mathrm{Lip}(\gamma)} + \|u\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}\|h\|_{\mathfrak{s}_0+3}^{\mathrm{Lip}(\gamma)}\big), \quad i = 1, 2. \tag{3.62}$$

(ii) The constant coefficients m_3, m_1 of $\mathcal{L}_5$ defined in (3.55) satisfy

$$|m_3 - 1| + |m_1| \le \varepsilon C, \tag{3.63}$$

$$|\partial_u m_3(u)[h]| + |\partial_u m_1(u)[h]| \le \varepsilon C \|h\|_\sigma. \tag{3.64}$$

Moreover, if $u(\lambda)$ is a Lipschitz family satisfying (3.61), then

$$|m_3 - 1|^{\mathrm{Lip}(\gamma)} + |m_1|^{\mathrm{Lip}(\gamma)} \le \varepsilon C. \tag{3.65}$$

(iii) The operator $\mathcal{R}$ defined in (3.55) satisfies:

$$|\mathcal{R}|_s \le \varepsilon C(s)(1 + \|u\|_{s+\sigma}), \tag{3.66}$$

$$|\partial_u \mathcal{R}(u)[h]|_s \le \varepsilon C(s)\big(\|h\|_{s+\sigma'} + \|u\|_{s+\sigma}\|h\|_{\mathfrak{s}_0+\sigma'}\big), \tag{3.67}$$

where $\sigma > \sigma'$ are defined in (3.58). Moreover, if $u(\lambda)$ is a Lipschitz family satisfying (3.61), then

$$|\mathcal{R}|_s^{\mathrm{Lip}(\gamma)} \le \varepsilon C(s)(1 + \|u\|_{s+\sigma}^{\mathrm{Lip}(\gamma)}). \tag{3.68}$$

Finally, in the reversible case, the maps Φ_i, Φ_i^{-1}, $i = 1, 2$ are reversibility preserving and $\mathcal{R}, \mathcal{L}_5 : X \to Y$ are reversible. In the Hamiltonian case the operator $\mathcal{L}_5$ is Hamiltonian.

PROOF. In section 7. ∎

Lemma 3.3. *In the same hypotheses of Lemma 3.2, for all $\varphi \in \mathbb{T}^\nu$, the operators $\mathcal{A}(\varphi), \mathcal{M}(\varphi), \mathcal{T}(\varphi), \mathcal{S}(\varphi)$ are invertible operators of the phase space $H_x^s := H^s(\mathbb{T})$, with*

$$\|\mathcal{A}^{\pm 1}(\varphi)h\|_{H_x^s} \le C(s)\big(\|h\|_{H_x^s} + \|u\|_{s+\mathfrak{s}_0+3}\|h\|_{H_x^1}\big), \tag{3.69}$$

$$\|(\mathcal{A}^{\pm 1}(\varphi) - I)h\|_{H_x^s} \le \varepsilon C(s)\big(\|h\|_{H_x^{s+1}} + \|u\|_{s+\mathfrak{s}_0+3}\|h\|_{H_x^2}\big), \tag{3.70}$$

$$\|(\mathcal{M}(\varphi)\mathcal{T}(\varphi)\mathcal{S}(\varphi))^{\pm 1}h\|_{H_x^s} \le C(s)\big(\|h\|_{H_x^s} + \|u\|_{s+\sigma}\|h\|_{H_x^1}\big), \tag{3.71}$$

$$\|((\mathcal{M}(\varphi)\mathcal{T}(\varphi)\mathcal{S}(\varphi))^{\pm 1} - I)h\|_{H_x^s} \le \varepsilon\gamma_0^{-1} C(s)\big(\|h\|_{H_x^{s+1}} + \|u\|_{s+\sigma}\|h\|_{H_x^1}\big). \tag{3.72}$$

PROOF. In section 7. ∎

4 Reduction of the linearized operator to constant coefficients

The goal of this section is to diagonalize the linear operator $\mathcal{L}_5$ obtained in (3.55), and therefore to complete the reduction of $\mathcal{L}$ in (3.1) into constant coefficients. For $\tau > \tau_0$ (see (1.2)) we define the constant

$$\beta := 7\tau + 6\,. \tag{4.1}$$

Theorem 4.1. *Let $f \in C^q$, see (1.3). Let $\gamma \in (0,1)$ and $\mathfrak{s}_0 \leq s \leq q - \sigma - \beta$ where σ is defined in (3.58), and β in (4.1). Let $u(\lambda)$ be a family of functions depending on the parameter $\lambda \in \Lambda_o \subset \Lambda := [1/2, 3/2]$ in a Lipschitz way, with*

$$\|u\|_{\mathfrak{s}_0+\sigma+\beta,\Lambda_o}^{\mathrm{Lip}(\gamma)} \leq 1\,. \tag{4.2}$$

Then there exist δ_0, C (depending on the data of the problem) such that, if

$$\varepsilon\gamma^{-1} \leq \delta_0\,, \tag{4.3}$$

then:

(i) **(Eigenvalues)** *$\forall \lambda \in \Lambda$ there exists a sequence*

$$\mu_j^\infty(\lambda) := \mu_j^\infty(\lambda, u) = \tilde{\mu}_j^0(\lambda) + r_j^\infty(\lambda)\,, \ \ \tilde{\mu}_j^0(\lambda) := \mathrm{i}\big(-\tilde{m}_3(\lambda)j^3 + \tilde{m}_1(\lambda)j\big)\,, \ j \in \mathbb{Z}\,, \tag{4.4}$$

where $\tilde{m}_3, \tilde{m}_1$ coincide with the coefficients of $\mathcal{L}_5$ in (3.55) for all $\lambda \in \Lambda_o$, and the corrections r_j^∞ satisfy

$$|\tilde{m}_3 - 1|^{\mathrm{Lip}(\gamma)} + |\tilde{m}_1|^{\mathrm{Lip}(\gamma)} + |r_j^\infty|_\Lambda^{\mathrm{Lip}(\gamma)} \leq \varepsilon C\,, \quad \forall j \in \mathbb{Z}\,. \tag{4.5}$$

Moreover, in the reversible case (i.e. (1.13) holds) or Hamiltonian case (i.e. (1.11) holds), all the eigenvalues μ_j^∞ are purely imaginary.

(ii) **(Conjugacy)**. *For all λ in*

$$\Lambda_\infty^{2\gamma} := \Lambda_\infty^{2\gamma}(u) := \Big\{\lambda \in \Lambda_o \ : \ |\mathrm{i}\lambda\bar{\omega} \cdot l + \mu_j^\infty(\lambda) - \mu_k^\infty(\lambda)| \geq 2\gamma|j^3 - k^3|\langle l\rangle^{-\tau}, \ \forall l \in \mathbb{Z}^\nu, \ j, k \in \mathbb{Z}\Big\} \tag{4.6}$$

there is a bounded, invertible linear operator $\Phi_\infty(\lambda) : H^s \to H^s$, with bounded inverse $\Phi_\infty^{-1}(\lambda)$, that conjugates $\mathcal{L}_5$ in (3.55) to constant coefficients, namely

$$\mathcal{L}_\infty(\lambda) := \Phi_\infty^{-1}(\lambda) \circ \mathcal{L}_5(\lambda) \circ \Phi_\infty(\lambda) = \lambda\bar{\omega} \cdot \partial_\varphi + \mathcal{D}_\infty(\lambda)\,, \quad \mathcal{D}_\infty(\lambda) := \mathrm{diag}_{j\in\mathbb{Z}}\mu_j^\infty(\lambda)\,. \tag{4.7}$$

The transformations $\Phi_\infty, \Phi_\infty^{-1}$ are close to the identity in matrix decay norm, with estimates

$$|\Phi_\infty(\lambda) - I|_{s,\Lambda_\infty^{2\gamma}}^{\mathrm{Lip}(\gamma)} + |\Phi_\infty^{-1}(\lambda) - I|_{s,\Lambda_\infty^{2\gamma}}^{\mathrm{Lip}(\gamma)} \leq \varepsilon\gamma^{-1}C(s)\big(1 + \|u\|_{s+\sigma+\beta,\Lambda_o}^{\mathrm{Lip}(\gamma)}\big)\,. \tag{4.8}$$

For all $\varphi \in \mathbb{T}^\nu$, the operator $\Phi_\infty(\varphi) : H_x^s \to H_x^s$ is invertible (where $H_x^s := H^s(\mathbb{T})$) with inverse $(\Phi_\infty(\varphi))^{-1} = \Phi_\infty^{-1}(\varphi)$, and

$$\|(\Phi_\infty^{\pm 1}(\varphi) - I)h\|_{H_x^s} \leq \varepsilon\gamma^{-1}C(s)\big(\|h\|_{H_x^s} + \|u\|_{s+\sigma+\beta+\mathfrak{s}_0}\|h\|_{H_x^1}\big)\,. \tag{4.9}$$

In the reversible case $\Phi_\infty, \Phi_\infty^{-1} : X \to X$, $Y \to Y$ are reversibility preserving, and $\mathcal{L}_\infty : X \to Y$ is reversible. In the Hamiltonian case the final $\mathcal{L}_\infty$ is Hamiltonian.

An important point of Theorem 4.1 is to require *only* the bound (4.2) for the low norm of u, but it provides the estimate for $\Phi_\infty^{\pm 1} - I$ in (4.8) also for the higher norms $|\cdot|_s$, depending also on the high norms of u. From Theorem 4.1 we shall deduce tame estimates for the inverse linearized operators in Theorem 4.3.

Note also that the set $\Lambda_\infty^{2\gamma}$ in (4.6) depends only of the final eigenvalues, and it is not defined inductively as in usual KAM theorems. This characterization of the set of parameters which fulfill all the required Melnikov non-resonance conditions (at any step of the iteration) was first observed in [6], [5] in an analytic setting. Theorem 4.1 extends this property also in a differentiable setting. A main advantage of this formulation is that it allows to discuss the measure estimates only once and not inductively: the Cantor set $\Lambda_\infty^{2\gamma}$ in (4.6)

could be empty (actually its measure $|\Lambda_\infty^{2\gamma}| = 1 - O(\gamma)$ as $\gamma \to 0$) but the functions $\mu_j^\infty(\lambda)$ are anyway well defined for all $\lambda \in \Lambda$, see (4.4). In particular we shall perform the measure estimates only along the nonlinear iteration, see section 5.

Theorem 4.1 is deduced from the following iterative Nash-Moser reducibility theorem for a linear operator of the form

$$\mathcal{L}_0 = \omega \cdot \partial_\varphi + \mathcal{D}_0 + \mathcal{R}_0 \,, \tag{4.10}$$

where $\omega = \lambda \bar{\omega}$,

$$\mathcal{D}_0 := m_3(\lambda, u(\lambda))\partial_{xxx} + m_1(\lambda, u(\lambda))\partial_x \,, \quad \mathcal{R}_0(\lambda, u(\lambda)) := \mathcal{R}(\lambda, u(\lambda)) \,, \tag{4.11}$$

the $m_3(\lambda, u(\lambda)), m_1(\lambda, u(\lambda)) \in \mathbb{R}$ and $u(\lambda)$ is defined for $\lambda \in \Lambda_o \subset \Lambda$. Clearly $\mathcal{L}_5$ in (3.55) has the form (4.10). Define

$$N_{-1} := 1 \,, \quad N_\nu := N_0^{\chi^\nu} \ \forall \nu \geq 0 \,, \quad \chi := 3/2 \tag{4.12}$$

(then $N_{\nu+1} = N_\nu^\chi, \forall \nu \geq 0$) and

$$\alpha := 7\tau + 4 \,, \quad \sigma_2 := \sigma + \beta \tag{4.13}$$

where σ is defined in (3.58) and β is defined in (4.1).

Theorem 4.2. (KAM reducibility) *Let $q > \sigma + \mathfrak{s}_0 + \beta$. There exist $C_0 > 0$, $N_0 \in \mathbb{N}$ large, such that, if*

$$N_0^{C_0} |\mathcal{R}_0|_{\mathfrak{s}_0+\beta}^{\mathrm{Lip}(\gamma)} \gamma^{-1} \leq 1, \tag{4.14}$$

then, for all $\nu \geq 0$:

(S1)$_\nu$ *There exists an operator*

$$\mathcal{L}_\nu := \omega \cdot \partial_\varphi + \mathcal{D}_\nu + \mathcal{R}_\nu \quad where \quad \mathcal{D}_\nu = \mathrm{diag}_{j \in \mathbb{Z}}\{\mu_j^\nu(\lambda)\} \tag{4.15}$$

$$\mu_j^\nu(\lambda) = \mu_j^0(\lambda) + r_j^\nu(\lambda), \quad \mu_j^0(\lambda) := -\mathrm{i}\big(m_3(\lambda, u(\lambda))j^3 - m_1(\lambda, u(\lambda))j\big), \quad j \in \mathbb{Z}, \tag{4.16}$$

defined for all $\lambda \in \Lambda_\nu^\gamma(u)$, where $\Lambda_0^\gamma(u) := \Lambda_o$ (is the domain of u), and, for $\nu \geq 1$,

$$\Lambda_\nu^\gamma := \Lambda_\nu^\gamma(u) := \Big\{\lambda \in \Lambda_{\nu-1}^\gamma : \big|\mathrm{i}\omega \cdot l + \mu_j^{\nu-1}(\lambda) - \mu_k^{\nu-1}(\lambda)\big| \geq \gamma \frac{|j^3 - k^3|}{\langle l \rangle^\tau} \ \forall |l| \leq N_{\nu-1}, \ j, k \in \mathbb{Z}\Big\}. \tag{4.17}$$

For $\nu \geq 0$, $r_j^\nu = \overline{r_{-j}^\nu}$, equivalently $\mu_j^\nu = \overline{\mu_{-j}^\nu}$, and

$$|r_j^\nu|^{\mathrm{Lip}(\gamma)} := |r_j^\nu|_{\Lambda_\nu^\gamma}^{\mathrm{Lip}(\gamma)} \leq \varepsilon C \,. \tag{4.18}$$

The remainder $\mathcal{R}_\nu$ is real (Definition 2.2) and, $\forall s \in [\mathfrak{s}_0, q - \sigma - \beta]$,

$$|\mathcal{R}_\nu|_s^{\mathrm{Lip}(\gamma)} \leq |\mathcal{R}_0|_{s+\beta}^{\mathrm{Lip}(\gamma)} N_{\nu-1}^{-\alpha} \,, \quad |\mathcal{R}_\nu|_{s+\beta}^{\mathrm{Lip}(\gamma)} \leq |\mathcal{R}_0|_{s+\beta}^{\mathrm{Lip}(\gamma)} N_{\nu-1} \,. \tag{4.19}$$

Moreover, for $\nu \geq 1$,

$$\mathcal{L}_\nu = \Phi_{\nu-1}^{-1} \mathcal{L}_{\nu-1} \Phi_{\nu-1} \,, \quad \Phi_{\nu-1} := I + \Psi_{\nu-1} \,, \tag{4.20}$$

where the map $\Psi_{\nu-1}$ is real, Töplitz in time $\Psi_{\nu-1} := \Psi_{\nu-1}(\varphi)$ (see (2.17)), and satisfies

$$|\Psi_{\nu-1}|_s^{\mathrm{Lip}(\gamma)} \leq |\mathcal{R}_0|_{s+\beta}^{\mathrm{Lip}(\gamma)} \gamma^{-1} N_{\nu-1}^{2\tau+1} N_{\nu-2}^{-\alpha} \,. \tag{4.21}$$

In the reversible case, $\mathcal{R}_\nu : X \to Y$, $\Psi_{\nu-1}, \Phi_{\nu-1}, \Phi_{\nu-1}^{-1}$ are reversibility preserving. Moreover, all the $\mu_j^\nu(\lambda)$ are purely imaginary and $\mu_j^\nu = -\mu_{-j}^\nu, \forall j \in \mathbb{Z}$.

(S2)$_\nu$ *For all $j \in \mathbb{Z}$, there exist Lipschitz extensions $\widetilde{\mu}_j^\nu(\cdot) : \Lambda \to \mathbb{R}$ of $\mu_j^\nu(\cdot) : \Lambda_\nu^\gamma \to \mathbb{R}$ satisfying, for $\nu \geq 1$,*

$$|\widetilde{\mu}_j^\nu - \widetilde{\mu}_j^{\nu-1}|^{\mathrm{Lip}(\gamma)} \leq |\mathcal{R}_{\nu-1}|_{\mathfrak{s}_0}^{\mathrm{Lip}(\gamma)} \,. \tag{4.22}$$

(S3)$_\nu$ *Let $u_1(\lambda)$, $u_2(\lambda)$, be Lipschitz families of Sobolev functions, defined for $\lambda \in \Lambda_o$ and such that conditions (4.2), (4.14) hold with $\mathcal{R}_0 := \mathcal{R}_0(u_i)$, $i = 1, 2$, see (4.11).*

Then, for $\nu \geq 0$, $\forall \lambda \in \Lambda_\nu^{\gamma_1}(u_1) \cap \Lambda_\nu^{\gamma_2}(u_2)$, with $\gamma_1, \gamma_2 \in [\gamma/2, 2\gamma]$,

$$|\mathcal{R}_\nu(u_2) - \mathcal{R}_\nu(u_1)|_{\mathfrak{s}_0} \leq \varepsilon N_{\nu-1}^{-\alpha} \|u_1 - u_2\|_{\mathfrak{s}_0 + \sigma_2}, \quad |\mathcal{R}_\nu(u_2) - \mathcal{R}_\nu(u_1)|_{\mathfrak{s}_0 + \beta} \leq \varepsilon N_{\nu-1} \|u_1 - u_2\|_{\mathfrak{s}_0 + \sigma_2}. \quad (4.23)$$

Moreover, for $\nu \geq 1$, $\forall s \in [\mathfrak{s}_0, \mathfrak{s}_0 + \beta]$, $\forall j \in \mathbb{Z}$,

$$\left| \left(r_j^\nu(u_2) - r_j^\nu(u_1) \right) - \left(r_j^{\nu-1}(u_2) - r_j^{\nu-1}(u_1) \right) \right| \leq |\mathcal{R}_{\nu-1}(u_2) - \mathcal{R}_{\nu-1}(u_1)|_{\mathfrak{s}_0}, \quad (4.24)$$

$$|r_j^\nu(u_2) - r_j^\nu(u_1)| \leq \varepsilon C \|u_1 - u_2\|_{\mathfrak{s}_0 + \sigma_2}. \quad (4.25)$$

(S4)$_\nu$ *Let u_1, u_2 like in **(S3)$_\nu$** and $0 < \rho < \gamma/2$. For all $\nu \geq 0$ such that*

$$\varepsilon C N_{\nu-1}^\tau \|u_1 - u_2\|_{\mathfrak{s}_0 + \sigma_2}^{\sup} \leq \rho \quad \Longrightarrow \quad \Lambda_\nu^\gamma(u_1) \subseteq \Lambda_\nu^{\gamma-\rho}(u_2). \quad (4.26)$$

Remark 4.1. *In the Hamiltonian case $\Psi_{\nu-1}$ is Hamiltonian and, instead of (4.20) we consider the symplectic map*

$$\Phi_{\nu-1} := \exp(\Psi_{\nu-1}). \quad (4.27)$$

The corresponding operators $\mathcal{L}_\nu$, $\mathcal{R}_\nu$ are Hamiltonian. Note that the operators (4.27) and (4.20) differ for an operator of order $\Psi_{\nu-1}^2$.

The proof of Theorem 4.2 is postponed in Subsection 4.1. We first give some consequences.

Corollary 4.1. (KAM transformation) $\forall \lambda \in \cap_{\nu \geq 0} \Lambda_\nu^\gamma$ *the sequence*

$$\widetilde{\Phi}_\nu := \Phi_0 \circ \Phi_1 \circ \cdots \circ \Phi_\nu \quad (4.28)$$

converges in $|\cdot|_s^{\mathrm{Lip}(\gamma)}$ to an operator Φ_∞ and

$$\left| \Phi_\infty - I \right|_s^{\mathrm{Lip}(\gamma)} + \left| \Phi_\infty^{-1} - I \right|_s^{\mathrm{Lip}(\gamma)} \leq C(s) \left| \mathcal{R}_0 \right|_{s+\beta}^{\mathrm{Lip}(\gamma)} \gamma^{-1}. \quad (4.29)$$

In the reversible case Φ_∞ and Φ_∞^{-1} are reversibility preserving.

PROOF. To simplify notations we write $|\cdot|_s$ for $|\cdot|_s^{\mathrm{Lip}(\gamma)}$. For all $\nu \geq 0$ we have $\widetilde{\Phi}_{\nu+1} = \widetilde{\Phi}_\nu \circ \Phi_{\nu+1} = \widetilde{\Phi}_\nu + \widetilde{\Phi}_\nu \Psi_{\nu+1}$ (see (4.20)) and so

$$|\widetilde{\Phi}_{\nu+1}|_{\mathfrak{s}_0} \overset{(2.9)}{\leq} |\widetilde{\Phi}_\nu|_{\mathfrak{s}_0} + C|\widetilde{\Phi}_\nu|_{\mathfrak{s}_0} |\Psi_{\nu+1}|_{\mathfrak{s}_0} \overset{(4.21)}{\leq} |\widetilde{\Phi}_\nu|_{\mathfrak{s}_0}(1 + \varepsilon_\nu) \quad (4.30)$$

where $\varepsilon_\nu := C' |\mathcal{R}_0|_{\mathfrak{s}_0+\beta}^{\mathrm{Lip}(\gamma)} \gamma^{-1} N_{\nu+1}^{2\tau+1} N_\nu^{-\alpha}$. Iterating (4.30) we get, for all ν,

$$|\widetilde{\Phi}_{\nu+1}|_{\mathfrak{s}_0} \leq |\widetilde{\Phi}_0|_{\mathfrak{s}_0} \Pi_{\nu \geq 0}(1 + \varepsilon_\nu) \leq |\Phi_0|_{\mathfrak{s}_0} e^{C|\mathcal{R}_0|_{\mathfrak{s}_0+\beta}^{\mathrm{Lip}(\gamma)} \gamma^{-1}} \leq 2 \quad (4.31)$$

using (4.21) (with $\nu = 1$, $s = \mathfrak{s}_0$) to estimate $|\Phi_0|_{\mathfrak{s}_0}$ and (4.14). The high norm of $\widetilde{\Phi}_{\nu+1} = \widetilde{\Phi}_\nu + \widetilde{\Phi}_\nu \Psi_{\nu+1}$ is estimated by (2.10), (4.31) (for $\widetilde{\Phi}_\nu$), as

$$
\begin{aligned}
|\widetilde{\Phi}_{\nu+1}|_s \quad &\leq \quad |\widetilde{\Phi}_\nu|_s (1 + C(s) |\Psi_{\nu+1}|_{\mathfrak{s}_0}) + C(s) |\Psi_{\nu+1}|_s \\
&\overset{(4.21),(4.13)}{\leq} \quad |\widetilde{\Phi}_\nu|_s (1 + \varepsilon_\nu^{(0)}) + \varepsilon_\nu^{(s)}, \quad \varepsilon_\nu^{(0)} := |\mathcal{R}_0|_{\mathfrak{s}_0+\beta} \gamma^{-1} N_\nu^{-1}, \quad \varepsilon_\nu^{(s)} := |\mathcal{R}_0|_{s+\beta} \gamma^{-1} N_\nu^{-1}.
\end{aligned}
$$

Iterating the above inequality and, using $\Pi_{j \geq 0}(1 + \varepsilon_j^{(0)}) \leq 2$, we get

$$|\widetilde{\Phi}_{\nu+1}|_s \leq_s \sum_{j=0}^\infty \varepsilon_j^{(s)} + |\widetilde{\Phi}_0|_s \leq C(s)\left(1 + |\mathcal{R}_0|_{s+\beta} \gamma^{-1}\right) \quad (4.32)$$

using $|\Phi_0|_s \leq 1 + C(s)|\mathcal{R}_0|_{s+\beta}\gamma^{-1}$. Finally, the $\widetilde{\Phi}_j$ a Cauchy sequence in norm $|\cdot|_s$ because

$$|\widetilde{\Phi}_{\nu+m} - \widetilde{\Phi}_\nu|_s \quad \leq \quad \sum_{j=\nu}^{\nu+m-1} |\widetilde{\Phi}_{j+1} - \widetilde{\Phi}_j|_s \overset{(2.10)}{\leq_s} \sum_{j=\nu}^{\nu+m-1} \left(|\widetilde{\Phi}_j|_s |\Psi_{j+1}|_{\mathfrak{s}_0} + |\widetilde{\Phi}_j|_{\mathfrak{s}_0} |\Psi_{j+1}|_s \right)$$
$$\overset{(4.32),(4.21),(4.31),(4.14)}{\leq_s} \sum_{j\geq\nu} |\mathcal{R}_0|_{s+\beta}\,\gamma^{-1} N_j^{-1} \leq_s |\mathcal{R}_0|_{s+\beta}\,\gamma^{-1} N_\nu^{-1}\,. \tag{4.33}$$

Hence $\widetilde{\Phi}_\nu \overset{|\cdot|_s}{\to} \Phi_\infty$. The bound for $\Phi_\infty - I$ in (4.29) follows by (4.33) with $m = \infty$, $\nu = 0$ and $|\widetilde{\Phi}_0 - I|_s = |\Psi_0|_s \ll \gamma^{-1}|\mathcal{R}_0|_{s+\beta}$. Then the estimate for $\Phi_\infty^{-1} - I$ follows by (2.13).

In the reversible case all the Φ_ν are reversibility preserving and so $\widetilde{\Phi}_\nu$, Φ_∞ are reversibility preserving. $\blacksquare$

Remark 4.2. *In the Hamiltonian case, the transformation $\widetilde{\Phi}_\nu$ in (4.28) is symplectic, because Φ_ν is symplectic for all ν (see Remark 4.1). Therefore Φ_∞ is also symplectic.*

Let us define for all $j \in \mathbb{Z}$

$$\mu_j^\infty(\lambda) = \lim_{\nu\to+\infty} \widetilde{\mu}_j^\nu(\lambda) = \tilde{\mu}_j^0 + r_j^\infty(\lambda), \quad r_j^\infty(\lambda) := \lim_{\nu\to+\infty} \tilde{r}_j^\nu(\lambda) \quad \forall\lambda \in \Lambda.$$

It could happen that $\Lambda_{\nu_0}^\gamma = \emptyset$ (see (4.17)) for some ν_0. In such a case the iterative process of Theorem 4.2 stops after finitely many steps. However, we can always set $\widetilde{\mu}_j^\nu := \widetilde{\mu}_j^{\nu_0}$, $\forall\nu \geq \nu_0$, and the functions $\mu_j^\infty : \Lambda \to \mathbb{R}$ are always well defined.

Corollary 4.2. (Final eigenvalues) *For all $\nu \in \mathbb{N}$, $j \in \mathbb{Z}$*

$$|\mu_j^\infty - \widetilde{\mu}_j^\nu|_\Lambda^{\mathrm{Lip}(\gamma)} = |r_j^\infty - \tilde{r}_j^\nu|_\Lambda^{\mathrm{Lip}(\gamma)} \leq C\,|\mathcal{R}_0|_{\mathfrak{s}_0+\beta}^{\mathrm{Lip}(\gamma)} N_{\nu-1}^{-\alpha}, \quad |\mu_j^\infty - \widetilde{\mu}_j^0|_\Lambda^{\mathrm{Lip}(\gamma)} = |r_j^\infty|_\Lambda^{\mathrm{Lip}(\gamma)} \leq C\,|\mathcal{R}_0|_{\mathfrak{s}_0+\beta}^{\mathrm{Lip}(\gamma)}\,. \tag{4.34}$$

PROOF. The bound (4.34) follows by (4.22) and (4.19) by summing the telescopic series. $\blacksquare$

Lemma 4.1. (Cantor set)
$$\Lambda_\infty^{2\gamma} \subset \cap_{\nu\geq0}\Lambda_\nu^\gamma. \tag{4.35}$$

PROOF. Let $\lambda \in \Lambda_\infty^{2\gamma}$. By definition $\Lambda_\infty^{2\gamma} \subset \Lambda_0^\gamma := \Lambda_o$. Then for all $\nu > 0$, $|l| \leq N_\nu$, $j \neq k$

$$\left|i\omega \cdot l + \mu_j^\nu - \mu_k^\nu\right| \quad \geq \quad \left|i\omega \cdot l + \mu_j^\infty - \mu_k^\infty\right| - \left|\mu_j^\nu - \mu_j^\infty\right| - \left|\mu_k^\nu - \mu_k^\infty\right|$$
$$\overset{(4.6),(4.34)}{\geq} \quad 2\gamma\left|j^3 - k^3\right| \langle l\rangle^{-\tau} - 2C|\mathcal{R}_0|_{\mathfrak{s}_0+\beta}N_{\nu-1}^{-\alpha} \geq \gamma\left|j^3 - k^3\right| \langle l\rangle^{-\tau}$$

because $\gamma|j^3 - k^3|\langle l\rangle^{-\tau} \geq \gamma N_\nu^{-\tau} \overset{(4.14)}{\geq} 2C|\mathcal{R}_0|_{\mathfrak{s}_0+\beta}N_{\nu-1}^{-\alpha}$. $\blacksquare$

Lemma 4.2. *For all $\lambda \in \Lambda_\infty^{2\gamma}(u)$,*

$$\mu_j^\infty(\lambda) = \overline{\mu_{-j}^\infty(\lambda)}, \quad r_j^\infty(\lambda) = \overline{r_{-j}^\infty(\lambda)}, \tag{4.36}$$

and in the reversible case

$$\mu_j^\infty(\lambda) = -\mu_{-j}^\infty(\lambda), \quad r_j^\infty(\lambda) = -r_{-j}^\infty(\lambda). \tag{4.37}$$

Actually in the reversible case $\mu_j^\infty(\lambda)$ are purely imaginary for all $\lambda \in \Lambda$.

PROOF. Formula (4.36) and (4.37) follow because, for all $\lambda \in \Lambda_\infty^{2\gamma} \subseteq \cap_{\nu\geq0}\Lambda_\nu^\gamma$ (see (4.35)), we have $\mu_j^\nu = \overline{\mu_{-j}^\nu}$, $r_j^\nu = \overline{r_{-j}^\nu}$, and, in the reversible case, the μ_j^ν are purely imaginary and $\mu_j^\nu = -\mu_{-j}^\nu$, $r_j^\nu = -r_{-j}^\nu$. The final statement follows because, in the reversible case, the $\mu_j^\nu(\lambda) \in i\mathbb{R}$ as well as its extension $\widetilde{\mu}_j^\nu(\lambda)$. $\blacksquare$

Remark 4.3. *In the reversible case, (4.37) imply that $\mu_0^\infty = r_0^\infty = 0$.*

Proof of Theorem 4.1. We apply Theorem 4.2 to the linear operator $\mathcal{L}_0 := \mathcal{L}_5$ in (3.55), where $\mathcal{R}_0 = \mathcal{R}$ defined in (4.11) satisfies

$$|\mathcal{R}_0|_{\mathfrak{s}_0+\beta}^{\mathrm{Lip}(\gamma)} \overset{(3.68)}{\leq} \varepsilon C(\mathfrak{s}_0+\beta)\left(1 + \|u\|_{\mathfrak{s}_0+\sigma+\beta}^{\mathrm{Lip}(\gamma)}\right) \overset{(4.2)}{\leq} 2\varepsilon C(\mathfrak{s}_0+\beta). \tag{4.38}$$

Then the smallness condition (4.14) is implied by (4.3) taking $\delta_0 := \delta_0(\nu)$ small enough.

For all $\lambda \in \Lambda_\infty^{2\gamma} \subset \cap_{\nu \geq 0}\Lambda_\nu^\gamma$ (see (4.35)), the operators

$$\mathcal{L}_\nu \overset{(4.15)}{=} \omega \cdot \partial_\varphi + \mathcal{D}_\nu + \mathcal{R}_\nu \overset{|\cdot|_s^{\mathrm{Lip}(\gamma)}}{\longrightarrow} \omega \cdot \partial_\varphi + \mathcal{D}_\infty =: \mathcal{L}_\infty, \quad \mathcal{D}_\infty := \mathrm{diag}_{j\in\mathbb{Z}}\mu_j^\infty \tag{4.39}$$

because

$$|\mathcal{D}_\nu - \mathcal{D}_\infty|_s^{\mathrm{Lip}(\gamma)} = \sup_{j\in\mathbb{Z}}\left|\mu_j^\nu - \mu_j^\infty\right|^{\mathrm{Lip}(\gamma)} \overset{(4.34)}{\leq} C\,|\mathcal{R}_0|_{\mathfrak{s}_0+\beta}^{\mathrm{Lip}(\gamma)}\,N_{\nu-1}^{-\alpha}, \quad |\mathcal{R}_\nu|_s^{\mathrm{Lip}(\gamma)} \overset{(4.19)}{\leq} |\mathcal{R}_0|_{s+\beta}^{\mathrm{Lip}(\gamma)}\,N_{\nu-1}^{-\alpha}.$$

Applying (4.20) iteratively we get $\mathcal{L}_\nu = \widetilde{\Phi}_{\nu-1}^{-1}\mathcal{L}_0\widetilde{\Phi}_{\nu-1}$ where $\widetilde{\Phi}_{\nu-1}$ is defined by (4.28) and $\widetilde{\Phi}_{\nu-1} \to \Phi_\infty$ in $|\;|_s$ (Corollary 4.1). Passing to the limit we deduce (4.7). Moreover (4.34) and (4.38) imply (4.5). Then (4.29), (3.68) (applied to $\mathcal{R}_0 = \mathcal{R}$) imply (4.8).

Estimate (4.9) follows from (2.12) (in $H_x^s(\mathbb{T})$), Lemma 2.4, and the bound (4.8).

In the reversible case, since Φ_∞, Φ_∞^{-1} are reversibility preserving (see Corollary 4.1), and $\mathcal{L}_0$ is reversible (see Remark 3.12 and Lemma 3.2), we get that $\mathcal{L}_\infty$ is reversible too. The eigenvalues μ_j^∞ are purely imaginary by Lemma 4.2.

In the Hamiltonian case, $\mathcal{L}_0 \equiv \mathcal{L}_5$ is Hamiltonian, Φ_∞ is symplectic, and therefore $\mathcal{L}_\infty = \Phi_\infty^{-1}\mathcal{L}_5\Phi_\infty$ (see (4.7)) is Hamiltonian, namely $\mathcal{D}_\infty$ has the structure $\mathcal{D}_\infty = \partial_x\mathcal{B}$, where $\mathcal{B} = \mathrm{diag}_{j\neq 0}\{b_j\}$ is self-adjoint. This means that $b_j \in \mathbb{R}$, and therefore $\mu_j^\infty = \mathrm{i}jb_j$ are all purely imaginary. $\blacksquare$

4.1 Proof of Theorem 4.2

PROOF OF $(\mathbf{Si})_0$, $i = 1, \ldots, 4$. Properties (4.15)-(4.19) in $(\mathbf{S1})_0$ hold by (4.10)-(4.11) with μ_j^0 defined in (4.16) and $r_j^0(\lambda) = 0$ (for (4.19) recall that $N_{-1} := 1$, see (4.12)). Moreover, since m_1, m_3 are real functions, μ_j^0 are purely imaginary, $\mu_j^0 = \overline{\mu_{-j}^0}$ and $\mu_j^0 = -\mu_{-j}^0$. In the reversible case, remark 3.12 implies that $\mathcal{R}_0 := \mathcal{R}$, $\mathcal{L}_0 := \mathcal{L}_5$ are reversible operators. Then there is nothing else to verify.

$(\mathbf{S2})_0$ holds extending from $\Lambda_0^\gamma := \Lambda_o$ to Λ the eigenvalues $\mu_j^0(\lambda)$, namely extending the functions $m_1(\lambda)$, $m_3(\lambda)$ to $\tilde{m}_1(\lambda)$, $\tilde{m}_3(\lambda)$, preserving the sup norm and the Lipschitz semi-norm, by Kirszbraun theorem.

$(\mathbf{S3})_0$ follows by (3.67), for $s = \mathfrak{s}_0, \mathfrak{s}_0 + \beta$, and (4.2), (4.13).

$(\mathbf{S4})_0$ is trivial because, by definition, $\Lambda_0^\gamma(u_1) = \Lambda_o = \Lambda_0^{\gamma-\rho}(u_2)$.

4.1.1 The reducibility step

We now describe the generic inductive step, showing how to define $\mathcal{L}_{\nu+1}$ (and Φ_ν, Ψ_ν, etc). To simplify notations, in this section we drop the index ν and we write $+$ for $\nu + 1$. We have

$$\begin{aligned}
\mathcal{L}\Phi h &= \omega \cdot \partial_\varphi(\Phi(h)) + \mathcal{D}\Phi h + \mathcal{R}\Phi h \\
&= \omega \cdot \partial_\varphi h + \Psi\omega \cdot \partial_\varphi h + (\omega \cdot \partial_\varphi\Psi)h + \mathcal{D}h + \mathcal{D}\Psi h + \mathcal{R}h + \mathcal{R}\Psi h \\
&= \Phi\left(\omega \cdot \partial_\varphi h + \mathcal{D}h\right) + \left(\omega \cdot \partial_\varphi\Psi + [\mathcal{D}, \Psi] + \Pi_N\mathcal{R}\right)h + \left(\Pi_N^\perp\mathcal{R} + \mathcal{R}\Psi\right)h \tag{4.40}
\end{aligned}$$

where $[\mathcal{D}, \Psi] := \mathcal{D}\Psi - \Psi\mathcal{D}$ and $\Pi_N\mathcal{R}$ is defined in (2.19).

Remark 4.4. *The application of the smoothing operator Π_N is necessary since we are performing a differentiable Nash-Moser scheme. Note also that Π_N regularizes only in time (see (2.19)) because the loss of derivatives of the inverse operator is only in φ (see (4.44) and the bound on the small divisors (4.17)).*

We look for a solution of the *homological* equation

$$\omega \cdot \partial_\varphi \Psi + [\mathcal{D}, \Psi] + \Pi_N \mathcal{R} = [\mathcal{R}] \qquad \text{where} \qquad [\mathcal{R}] := \mathrm{diag}_{j \in \mathbb{Z}} \mathcal{R}_j^j(0) \,. \tag{4.41}$$

Lemma 4.3. (Homological equation) *For all* $\lambda \in \Lambda_{\nu+1}^\gamma$, *(see* (4.17)*) there exists a unique solution* $\Psi := \Psi(\varphi)$ *of the homological equation* (4.41)*. The map* Ψ *satisfies*

$$|\Psi|_s^{\mathrm{Lip}(\gamma)} \le C N^{2\tau+1} \gamma^{-1} |\mathcal{R}|_s^{\mathrm{Lip}(\gamma)} \,. \tag{4.42}$$

Moreover if $\gamma/2 \le \gamma_1, \gamma_2 \le 2\gamma$ *and if* $u_1(\lambda)$, $u_2(\lambda)$ *are Lipschitz functions, then* $\forall s \in [\mathfrak{s}_0, \mathfrak{s}_0 + \beta]$, $\lambda \in \Lambda_{\nu+1}^{\gamma_1}(u_1) \cap \Lambda_{\nu+1}^{\gamma_2}(u_2)$

$$|\Delta_{12}\Psi|_s \le C N^{2\tau+1} \gamma^{-1} \Big(|\mathcal{R}(u_2)|_s \|u_1 - u_2\|_{\mathfrak{s}_0 + \sigma_2} + |\Delta_{12}\mathcal{R}|_s \Big) \tag{4.43}$$

where we define $\Delta_{12}\Psi := \Psi(u_1) - \Psi(u_2)$.
In the reversible case, Ψ *is reversibility-preserving.*

PROOF. Since $\mathcal{D} := \mathrm{diag}_{j \in \mathbb{Z}}(\mu_j)$ we have $[\mathcal{D}, \Psi]_j^k = (\mu_j - \mu_k)\Psi_j^k(\varphi)$ and (4.41) amounts to

$$\omega \cdot \partial_\varphi \Psi_j^k(\varphi) + (\mu_j - \mu_k)\Psi_j^k(\varphi) + \mathcal{R}_j^k(\varphi) = [\mathcal{R}]_j^k \,, \quad \forall j, k \in \mathbb{Z} \,,$$

whose solutions are $\Psi_j^k(\varphi) = \sum_{l \in \mathbb{Z}^\nu} \Psi_j^k(l) e^{il \cdot \varphi}$ with coefficients

$$\Psi_j^k(l) := \begin{cases} \dfrac{\mathcal{R}_j^k(l)}{\delta_{ljk}(\lambda)} & \text{if } (j - k, l) \ne (0, 0) \text{ and } |l| \le N \,, \text{ where } \delta_{ljk}(\lambda) := i\omega \cdot l + \mu_j - \mu_k, \\ 0 & \text{otherwise.} \end{cases} \tag{4.44}$$

Note that, for all $\lambda \in \Lambda_{\nu+1}^\gamma$, by (4.17) and (1.2), if $j \ne k$ or $l \ne 0$ the divisors $\delta_{ljk}(\lambda) \ne 0$. Recalling the definition of the s-norm in (2.3) we deduce by (4.44), (4.17), (1.2), that

$$|\Psi|_s \le \gamma^{-1} N^\tau |\mathcal{R}|_s \,, \quad \forall \lambda \in \Lambda_{\nu+1}^\gamma \,. \tag{4.45}$$

For $\lambda_1, \lambda_2 \in \Lambda_{\nu+1}^\gamma$,

$$|\Psi_j^k(l)(\lambda_1) - \Psi_j^k(l)(\lambda_2)| \le \frac{|\mathcal{R}_j^k(l)(\lambda_1) - \mathcal{R}_j^k(l)(\lambda_2)|}{|\delta_{ljk}(\lambda_1)|} + |\mathcal{R}_j^k(l)(\lambda_2)| \frac{|\delta_{ljk}(\lambda_1) - \delta_{ljk}(\lambda_2)|}{|\delta_{ljk}(\lambda_1)||\delta_{ljk}(\lambda_2)|} \tag{4.46}$$

and, since $\omega = \lambda\bar\omega$,

$$|\delta_{ljk}(\lambda_1) - \delta_{ljk}(\lambda_2)| \overset{(4.44)}{=} |(\lambda_1 - \lambda_2)\bar\omega \cdot l + (\mu_j - \mu_k)(\lambda_1) - (\mu_j - \mu_k)(\lambda_2)| \tag{4.47}$$

$$\overset{(4.16)}{\le} |\lambda_1 - \lambda_2||\bar\omega \cdot l| + |m_3(\lambda_1) - m_3(\lambda_2)||j^3 - k^3| + |m_1(\lambda_1) - m_1(\lambda_2)||j - k|$$
$$\qquad + |r_j(\lambda_1) - r_j(\lambda_2)| + |r_k(\lambda_1) - r_k(\lambda_2)|$$
$$\ll |\lambda_1 - \lambda_2| \Big(|l| + \varepsilon\gamma^{-1}|j^3 - k^3| + \varepsilon\gamma^{-1}|j - k| + \varepsilon\gamma^{-1} \Big) \tag{4.48}$$

because

$$\gamma|m_3|^{\mathrm{lip}} = \gamma|m_3 - 1|^{\mathrm{lip}} \le |m_3 - 1|^{\mathrm{Lip}(\gamma)} \le \varepsilon C, \quad |m_1|^{\mathrm{Lip}(\gamma)} \le \varepsilon C, \quad |r_j|^{\mathrm{Lip}(\gamma)} \le \varepsilon C \quad \forall j \in \mathbb{Z}.$$

Hence, for $j \ne k$, $\varepsilon\gamma^{-1} \le 1$,

$$\frac{|\delta_{ljk}(\lambda_1) - \delta_{ljk}(\lambda_2)|}{|\delta_{ljk}(\lambda_1)||\delta_{ljk}(\lambda_2)|} \overset{(4.48),(4.17)}{\ll} |\lambda_1 - \lambda_2| \Big(|l| + |j^3 - k^3| \Big) \frac{\langle l \rangle^{2\tau}}{\gamma^2 |j^3 - k^3|^2} \ll |\lambda_1 - \lambda_2| N^{2\tau+1} \gamma^{-2} \tag{4.49}$$

26

for $|l| \leq N$. Finally, recalling (2.3), the bounds (4.46), (4.49) and (4.45) imply (4.42). Now we prove (4.43). By (4.44), for any $\lambda \in \Lambda_{\nu+1}^{\gamma_1}(u_1) \cap \Lambda_{\nu+1}^{\gamma_2}(u_2)$, $l \in \mathbb{Z}^\nu$, $j \neq k$, we get

$$\Delta_{12}\Psi_j^k(l) = \frac{\Delta_{12}\mathcal{R}_j^k(l)}{\delta_{ljk}(u_1)} - \mathcal{R}_j^k(l)(u_2)\frac{\Delta_{12}\delta_{ljk}}{\delta_{ljk}(u_1)\delta_{ljk}(u_2)} \tag{4.50}$$

where

$$|\Delta_{12}\delta_{ljk}| = |\Delta_{12}(\mu_j - \mu_k)| \leq |\Delta_{12}m_3||j^3 - k^3| + |\Delta_{12}m_1||j - k| + |\Delta_{12}r_j| + |\Delta_{12}r_k|$$
$$\underset{(3.64),(4.25)}{\ll} \varepsilon|j^3 - k^3|\|u_1 - u_2\|_{\mathfrak{s}_0+\sigma_2}. \tag{4.51}$$

Then (4.50), (4.51), $\varepsilon\gamma^{-1} \leq 1$, $\gamma_1^{-1}, \gamma_2^{-1} \leq \gamma^{-1}$ imply

$$|\Delta_{12}\Psi_j^k(l)| \ll N^{2\tau}\gamma^{-1}\Big(|\Delta_{12}\mathcal{R}_j^k(l)| + |\mathcal{R}_j^k(l)(u_2)|\|u_1 - u_2\|_{\mathfrak{s}_0+\sigma_2}\Big)$$

and so (4.43) (in fact, (4.43) holds with 2τ instead of $2\tau + 1$).

In the reversible case $i\omega \cdot l + \mu_j - \mu_k \in i\mathbb{R}$, $\overline{\mu_{-j}} = \mu_j$ and $\mu_{-j} = -\mu_j$. Hence Lemma 2.6 and (4.44) imply

$$\overline{\Psi_{-j}^{-k}(-l)} = \frac{\overline{\mathcal{R}_{-j}^{-k}(-l)}}{-i\omega \cdot (-l) + \overline{\mu_{-j}} - \overline{\mu_{-k}}} = \frac{\mathcal{R}_j^k(l)}{i\omega \cdot l + \mu_j - \mu_k} = \Psi_j^k(l)$$

and so Ψ is real, again by Lemma 2.6. Moreover, since $\mathcal{R} : X \to Y$,

$$\Psi_{-j}^{-k}(-l) = \frac{\mathcal{R}_{-j}^{-k}(-l)}{i\omega \cdot (-l) + \mu_{-j} - \mu_{-k}} = \frac{-\mathcal{R}_j^k(l)}{i\omega \cdot (-l) - \mu_j + \mu_k} = \Psi_j^k(l)$$

which implies $\Psi : X \to X$ by Lemma 2.6. Similarly we get $\Psi : Y \to Y$. ∎

Remark 4.5. *In the Hamiltonian case $\mathcal{R}$ is Hamiltonian and the solution Ψ in (4.44) of the homological equation is Hamiltonian, because $\overline{\delta_{l,j,k}} = \delta_{-l,k,j}$ and, in terms of matrix elements, an operator $G(\varphi)$ is self-adjoint if and only if $\overline{G_j^k(l)} = G_k^j(-l)$.*

Let Ψ be the solution of the homological equation (4.41) which has been constructed in Lemma 4.3. By Lemma 2.3, if $C(\mathfrak{s}_0)|\Psi|_{\mathfrak{s}_0} < 1/2$ then $\Phi := I + \Psi$ is invertible and by (4.40) (and (4.41)) we deduce that

$$\mathcal{L}_+ := \Phi^{-1}\mathcal{L}\Phi = \omega \cdot \partial_\varphi + \mathcal{D}_+ + \mathcal{R}_+, \tag{4.52}$$

where

$$\mathcal{D}_+ := \mathcal{D} + [\mathcal{R}], \quad \mathcal{R}_+ := \Phi^{-1}\Big(\Pi_N^\perp \mathcal{R} + \mathcal{R}\Psi - \Psi[\mathcal{R}]\Big). \tag{4.53}$$

Note that $\mathcal{L}_+$ has the same form of $\mathcal{L}$, but the remainder $\mathcal{R}_+$ is the sum of a quadratic function of $\Psi, \mathcal{R}$ and a remainder supported on high modes.

Lemma 4.4. (New diagonal part). *The eigenvalues of*

$$\mathcal{D}_+ = \text{diag}_{j\in\mathbb{Z}}\{\mu_j^+(\lambda)\}, \quad where \quad \mu_j^+ := \mu_j + \mathcal{R}_j^j(0) = \mu_j^0 + r_j + \mathcal{R}_j^j(0) = \mu_j^0 + r_j^+, \quad r_j^+ := r_j + \mathcal{R}_j^j(0),$$

satisfy $\mu_j^+ = \overline{\mu_{-j}^+}$ and

$$|\mu_j^+ - \mu_j|^{\text{lip}} = |r_j^+ - r_j|^{\text{lip}} = |\mathcal{R}_j^j(0)|^{\text{lip}} \leq |\mathcal{R}|_{\mathfrak{s}_0}^{\text{lip}}, \quad \forall j \in \mathbb{Z}. \tag{4.54}$$

Moreover if $u_1(\lambda)$, $u_2(\lambda)$ are Lipschitz functions, then for all $\lambda \in \Lambda_\nu^{\gamma_1}(u_1) \cap \Lambda_\nu^{\gamma_2}(u_2)$

$$|\Delta_{12}r_j^+ - \Delta_{12}r_j| \leq |\Delta_{12}\mathcal{R}|_{\mathfrak{s}_0}. \tag{4.55}$$

In the reversible case, all the μ_j^+ are purely imaginary and satisfy $\mu_j^+ = -\mu_{-j}^+$ for all $j \in \mathbb{Z}$.

PROOF. The estimates (4.54)-(4.55) follow using (2.4) because $|\mathcal{R}_j^j(0)|^{\mathrm{lip}} = |\mathcal{R}_{(l,j)}^{(l,j)}|^{\mathrm{lip}} \leq |\mathcal{R}|_0^{\mathrm{lip}} \leq |\mathcal{R}|_{\mathfrak{s}_0}^{\mathrm{lip}}$ and

$$|\Delta_{12}r_j^+ - \Delta_{12}r_j| = |\Delta_{12}\mathcal{R}_j^j(0)| = |\Delta_{12}\mathcal{R}_{(l,j)}^{(l,j)}| \leq |\Delta_{12}\mathcal{R}|_0 \leq |\Delta_{12}\mathcal{R}|_{\mathfrak{s}_0}.$$

Since $\mathcal{R}$ is real, by Lemma 2.6,

$$\mathcal{R}_j^k(l) = \overline{\mathcal{R}_{-j}^{-k}(-l)} \qquad \Longrightarrow \qquad \mathcal{R}_j^j(0) = \overline{\mathcal{R}_{-j}^{-j}(0)}$$

and so $\mu_j^+ = \overline{\mu_{-j}^+}$. If $\mathcal{R}$ is also reversible, by Lemma 2.6,

$$\mathcal{R}_j^k(l) = -\mathcal{R}_{-j}^{-k}(-l), \quad \mathcal{R}_j^k(l) = \overline{\mathcal{R}_{-j}^{-k}(-l)} = -\overline{\mathcal{R}_j^k(l)}.$$

We deduce that $\mathcal{R}_j^j(0) = -\mathcal{R}_{-j}^{-j}(0)$, $\mathcal{R}_j^j(0) \in i\mathbb{R}$ and therefore, $\mu_j^+ = -\mu_{-j}^+$ and $\mu_j^+ \in i\mathbb{R}$. ∎

Remark 4.6. *In the Hamiltonian case, $\mathcal{D}_\nu$ is Hamiltonian, namely $\mathcal{D}_\nu = \partial_x \mathcal{B}$ where $\mathcal{B} = \mathrm{diag}_{j \neq 0}\{b_j\}$ is self-adjoint. This means that $b_j \in \mathbb{R}$, and therefore all $\mu_j^\nu = ijb_j$ are purely imaginary.*

4.1.2 The iteration

Let $\nu \geq 0$, and suppose that the statements $(\mathbf{Si})_\nu$ are true. We prove $(\mathbf{Si})_{\nu+1}$, $i = 1, \ldots, 4$. To simplify notations we write $|\cdot|_s$ instead of $|\cdot|_s^{\mathrm{Lip}(\gamma)}$.

PROOF OF $(\mathbf{S1})_{\nu+1}$. By $(\mathbf{S1})_\nu$, the eigenvalues μ_j^ν are defined on Λ_ν^γ. Therefore the set $\Lambda_{\nu+1}^\gamma$ is well-defined. By Lemma 4.3, for all $\lambda \in \Lambda_{\nu+1}^\gamma$ there exists a real solution Ψ_ν of the homological equation (4.41) which satisfies, $\forall s \in [\mathfrak{s}_0, q - \sigma - \beta]$,

$$|\Psi_\nu|_s \overset{(4.42)}{\prec} N_\nu^{2\tau+1} |\mathcal{R}_\nu|_s \gamma^{-1} \overset{(4.19)}{\prec} |\mathcal{R}_0|_{s+\beta} \gamma^{-1} N_\nu^{2\tau+1} N_{\nu-1}^{-\alpha} \tag{4.56}$$

which is (4.21) at the step $\nu + 1$. In particular, for $s = \mathfrak{s}_0$,

$$C(\mathfrak{s}_0) |\Psi_\nu|_{\mathfrak{s}_0} \overset{(4.56)}{\leq} C(\mathfrak{s}_0) |\mathcal{R}_0|_{\mathfrak{s}_0+\beta} \gamma^{-1} N_\nu^{2\tau+1} N_{\nu-1}^{-\alpha} \overset{(4.14)}{\leq} 1/2 \tag{4.57}$$

for N_0 large enough. Then the map $\Phi_\nu := I + \Psi_\nu$ is invertible and, by (2.13),

$$\left|\Phi_\nu^{-1}\right|_{\mathfrak{s}_0} \leq 2, \quad \left|\Phi_\nu^{-1}\right|_s \leq 1 + C(s)|\Psi_\nu|_s. \tag{4.58}$$

Hence (4.52)-(4.53) imply $\mathcal{L}_{\nu+1} := \Phi_\nu^{-1}\mathcal{L}_\nu\Phi_\nu = \omega \cdot \partial_\varphi + \mathcal{D}_{\nu+1} + \mathcal{R}_{\nu+1}$ where (see Lemma 4.4)

$$\mathcal{D}_{\nu+1} := \mathcal{D}_\nu + [\mathcal{R}_\nu] = \mathrm{diag}_{j \in \mathbb{Z}}(\mu_j^{\nu+1}), \quad \mu_j^{\nu+1} := \mu_j^\nu + (\mathcal{R}_\nu)_j^j(0), \tag{4.59}$$

with $\mu_j^{\nu+1} = \overline{\mu_{-j}^{\nu+1}}$ and

$$\mathcal{R}_{\nu+1} := \Phi_\nu^{-1} H_\nu, \quad H_\nu := \Pi_{N_\nu}^\perp \mathcal{R}_\nu + \mathcal{R}_\nu \Psi_\nu - \Psi_\nu[\mathcal{R}_\nu]. \tag{4.60}$$

In the reversible case, $\mathcal{R}_\nu : X \to Y$, therefore, by Lemma 4.3, Ψ_ν, Φ_ν, Φ_ν^{-1} are reversibility preserving, and then, by formula (4.60), also $\mathcal{R}_{\nu+1} : X \to Y$.

Let us prove the estimates (4.19) for $\mathcal{R}_{\nu+1}$. For all $s \in [\mathfrak{s}_0, q - \sigma - \beta]$ we have

$$|\mathcal{R}_{\nu+1}|_s \overset{(4.60),(2.10)}{\leq_s} |\Phi_\nu^{-1}|_{\mathfrak{s}_0}\left(|\Pi_{N_\nu}^\perp \mathcal{R}_\nu|_s + |\mathcal{R}_\nu|_s|\Psi_\nu|_{\mathfrak{s}_0} + |\mathcal{R}_\nu|_{\mathfrak{s}_0}|\Psi_\nu|_s\right) + |\Phi_\nu^{-1}|_s\left(|\Pi_{N_\nu}^\perp \mathcal{R}_\nu|_{\mathfrak{s}_0} + |\mathcal{R}_\nu|_{\mathfrak{s}_0}|\Psi_\nu|_{\mathfrak{s}_0}\right)$$

$$\overset{(4.58)}{\leq_s} 2\left(|\Pi_{N_\nu}^\perp \mathcal{R}_\nu|_s + |\mathcal{R}_\nu|_s|\Psi_\nu|_{\mathfrak{s}_0} + |\mathcal{R}_\nu|_{\mathfrak{s}_0}|\Psi_\nu|_s\right) + (1 + |\Psi_\nu|_s)\left(|\Pi_{N_\nu}^\perp \mathcal{R}_\nu|_{\mathfrak{s}_0} + |\mathcal{R}_\nu|_{\mathfrak{s}_0}|\Psi_\nu|_{\mathfrak{s}_0}\right)$$

$$\overset{(4.57)}{\leq_s} |\Pi_{N_\nu}^\perp \mathcal{R}_\nu|_s + |\mathcal{R}_\nu|_s|\Psi_\nu|_{\mathfrak{s}_0} + |\mathcal{R}_\nu|_{\mathfrak{s}_0}|\Psi_\nu|_s \overset{(4.42)}{\leq_s} |\Pi_{N_\nu}^\perp \mathcal{R}_\nu|_s + N_\nu^{2\tau+1}\gamma^{-1}|\mathcal{R}_\nu|_s|\mathcal{R}_\nu|_{\mathfrak{s}_0}. \tag{4.61}$$

Hence (4.61) and (2.20) imply

$$|\mathcal{R}_{\nu+1}|_s \leq_s N_\nu^{-\beta}|\mathcal{R}_\nu|_{s+\beta} + N_\nu^{2\tau+1}\gamma^{-1}|\mathcal{R}_\nu|_s|\mathcal{R}_\nu|_{\mathfrak{s}_0} \tag{4.62}$$

which shows that the iterative scheme is quadratic plus a super-exponentially small term. In particular

$$|\mathcal{R}_{\nu+1}|_s \overset{(4.62),(4.19)}{\leq_s} N_\nu^{-\beta}|\mathcal{R}_0|_{s+\beta}N_{\nu-1} + N_\nu^{2\tau+1}\gamma^{-1}|\mathcal{R}_0|_{s+\beta}|\mathcal{R}_0|_{\mathfrak{s}_0+\beta}N_{\nu-1}^{-2\alpha} \overset{(4.1),(4.13),(4.14)}{\leq} |\mathcal{R}_0|_{s+\beta}N_\nu^{-\alpha}$$

($\chi = 3/2$) which is the first inequality of (4.19) at the step $\nu + 1$. The next key step is to control the divergence of the high norm $|\mathcal{R}_{\nu+1}|_{s+\beta}$. By (4.61) (with $s + \beta$ instead of s) we get

$$|\mathcal{R}_{\nu+1}|_{s+\beta} \leq_{s+\beta} |\mathcal{R}_\nu|_{s+\beta} + N_\nu^{2\tau+1}\gamma^{-1}|\mathcal{R}_\nu|_{s+\beta}|\mathcal{R}_\nu|_{\mathfrak{s}_0} \tag{4.63}$$

(the difference with respect to (4.62) is that we do not apply to $|\Pi_{N_\nu}^\perp \mathcal{R}_\nu|_{s+\beta}$ any smoothing). Then (4.63), (4.19), (4.14), (4.13) imply the inequality

$$|\mathcal{R}_{\nu+1}|_{s+\beta} \leq C(s+\beta)|\mathcal{R}_\nu|_{s+\beta},$$

whence, iterating,

$$|\mathcal{R}_{\nu+1}|_{s+\beta} \leq N_\nu|\mathcal{R}_0|_{s+\beta}$$

for $N_0 := N_0(s,\beta)$ large enough, which is the second inequality of (4.19) with index $\nu + 1$.

By Lemma 4.4 the eigenvalues $\mu_j^{\nu+1} := \mu_j^0 + r_j^{\nu+1}$, defined on $\Lambda_{\nu+1}^\gamma$, satisfy $\mu_j^{\nu+1} = \overline{\mu_{-j}^{\nu+1}}$, and, in the reversible case, the $\mu_j^{\nu+1}$ are purely imaginary and $\mu_j^{\nu+1} = -\mu_{-j}^{\nu+1}$.

It remains only to prove (4.18) for $\nu + 1$, which is proved below.

PROOF OF $(\mathbf{S2})_{\nu+1}$. By (4.54),

$$|\mu_j^{\nu+1} - \mu_j^\nu|^{\text{Lip}(\gamma)} = |r_j^{\nu+1} - r_j^\nu|^{\text{Lip}(\gamma)} \leq |\mathcal{R}_\nu|_{\mathfrak{s}_0}^{\text{Lip}(\gamma)} \overset{(4.19)}{\leq} |\mathcal{R}_0|_{\mathfrak{s}_0+\beta}^{\text{Lip}(\gamma)} N_{\nu-1}^{-\alpha}. \tag{4.64}$$

By Kirszbraun theorem, we extend the function $\mu_j^{\nu+1} - \mu_j^\nu = r_j^{\nu+1} - r_j^\nu$ to the whole Λ, still satisfying (4.64). In this way we define $\tilde{\mu}_j^{\nu+1}$. Finally (4.18) follows summing all the terms in (4.64) and using (3.68).

PROOF OF $(\mathbf{S3})_{\nu+1}$. Set, for brevity,

$$\mathcal{R}_\nu^i := \mathcal{R}_\nu(u_i), \quad \Psi_{\nu-1}^i := \Psi_{\nu-1}(u_i), \quad \Phi_{\nu-1}^i := \Phi_{\nu-1}(u_i), \quad H_{\nu-1}^i := H_{\nu-1}(u_i), \quad i := 1,2,$$

which are all operators defined for $\lambda \in \Lambda_\nu^{\gamma_1}(u_1) \cap \Lambda_\nu^{\gamma_2}(u_2)$. By Lemma 4.3 one can construct $\Psi_\nu^i := \Psi_\nu(u_i)$, $\Phi_\nu^i := \Phi_\nu(u_i)$, $i = 1,2$, for all $\lambda \in \Lambda_{\nu+1}^{\gamma_1}(u_1) \cap \Lambda_{\nu+1}^{\gamma_2}(u_2)$. One has

$$|\Delta_{12}\Psi_\nu|_{\mathfrak{s}_0} \overset{(4.43)}{\prec} N_\nu^{2\tau+1}\gamma^{-1}\Big(|\mathcal{R}_\nu(u_2)|_{\mathfrak{s}_0}\|u_2 - u_1\|_{\mathfrak{s}_0+\sigma_2} + |\Delta_{12}\mathcal{R}_\nu|_{\mathfrak{s}_0}\Big)$$

$$\overset{(4.19),(4.23)}{\prec} N_\nu^{2\tau+1}N_{\nu-1}^{-\alpha}\gamma^{-1}\big(|\mathcal{R}_0|_{\mathfrak{s}_0+\beta} + \varepsilon\big)\|u_2 - u_1\|_{\mathfrak{s}_0+\sigma_2}$$

$$\overset{(3.68),(4.2)}{\prec} N_\nu^{2\tau+1}N_{\nu-1}^{-\alpha}\varepsilon\gamma^{-1}\|u_2 - u_1\|_{\mathfrak{s}_0+\sigma_2} \leq \|u_2 - u_1\|_{\mathfrak{s}_0+\sigma_2}. \tag{4.65}$$

for $\varepsilon\gamma^{-1}$ small (and (4.13)). By (2.14), applied to $\Phi := \Phi_\nu$, and (4.65), we get

$$|\Delta_{12}\Phi_\nu^{-1}|_s \leq_s \big(|\Psi_\nu^1|_s + |\Psi_\nu^2|_s\big)\|u_1 - u_2\|_{\mathfrak{s}_0+\sigma_2} + |\Delta_{12}\Psi_\nu|_s \tag{4.66}$$

which implies for $s = \mathfrak{s}_0$, and using (4.21), (4.14), (4.65)

$$|\Delta_{12}\Phi_\nu^{-1}|_{\mathfrak{s}_0} \prec \|u_1 - u_2\|_{\mathfrak{s}_0+\sigma_2}. \tag{4.67}$$

Let us prove the estimates (4.23) for $\Delta_{12}\mathcal{R}_{\nu+1}$, which is defined on $\lambda \in \Lambda_{\nu+1}^{\gamma_1}(u_1) \cap \Lambda_{\nu+1}^{\gamma_2}(u_2)$. For all $s \in [\mathfrak{s}_0, \mathfrak{s}_0+\beta]$, using the interpolation (2.7) and (4.60),

$$|\Delta_{12}\mathcal{R}_{\nu+1}|_s \leq_s |\Delta_{12}\Phi_\nu^{-1}|_s|H_\nu^1|_{\mathfrak{s}_0} + |\Delta_{12}\Phi_\nu^{-1}|_{\mathfrak{s}_0}|H_\nu^1|_s + |(\Phi_\nu^2)^{-1}|_s|\Delta_{12}H_\nu|_{\mathfrak{s}_0} + |(\Phi_\nu^2)^{-1}|_{\mathfrak{s}_0}|\Delta_{12}H_\nu|_s. \tag{4.68}$$

We estimate the above terms separately. Set for brevity $A_s^\nu := |\mathcal{R}_\nu(u_1)|_s + |\mathcal{R}_\nu(u_2)|_s$. By (4.60) and (2.7),

$$
|\Delta_{12} H_\nu|_s \quad \leq_s \quad \left|\Pi_{N_\nu}^\perp \Delta_{12}\mathcal{R}_\nu\right|_s + |\Delta_{12}\Psi_\nu|_s |\mathcal{R}_\nu^1|_{\mathfrak{s}_0} + |\Delta_{12}\Psi_\nu|_{\mathfrak{s}_0} |\mathcal{R}_\nu^1|_s + |\Psi_\nu^2|_s |\Delta_{12}\mathcal{R}_\nu|_{\mathfrak{s}_0} + |\Psi_\nu^2|_{\mathfrak{s}_0} |\Delta_{12}\mathcal{R}_\nu|_s
$$

$$
\overset{(4.42),(4.43)}{\leq_s} \quad \left|\Pi_{N_\nu}^\perp \Delta_{12}\mathcal{R}_\nu\right|_s + N_\nu^{2\tau+1}\gamma^{-1} A_{\mathfrak{s}_0}^\nu A_s^\nu \|u_1 - u_2\|_{\mathfrak{s}_0+\sigma_2}
$$
$$
+ N_\nu^{2\tau+1}\gamma^{-1} A_s^\nu |\Delta_{12}\mathcal{R}_\nu|_{\mathfrak{s}_0} + N_\nu^{2\tau+1}\gamma^{-1} A_{\mathfrak{s}_0}^\nu |\Delta_{12}\mathcal{R}_\nu|_s \,. \tag{4.69}
$$

Estimating the four terms in the right hand side of (4.68) in the same way, using (4.66), (4.60), (4.42), (4.43), (4.21), (4.67), (4.58), (4.69), (4.19), we deduce

$$
|\Delta_{12}\mathcal{R}_{\nu+1}|_s \quad \leq_s \quad |\Pi_{N_\nu}^\perp \Delta_{12}\mathcal{R}_\nu|_s + N_\nu^{2\tau+1}\gamma^{-1} A_s^\nu A_{\mathfrak{s}_0}^\nu \|u_1 - u_2\|_{\mathfrak{s}_0+\sigma_2}
$$
$$
+ N_\nu^{2\tau+1}\gamma^{-1} A_s^\nu |\Delta_{12}\mathcal{R}_\nu|_{\mathfrak{s}_0} + N_\nu^{2\tau+1}\gamma^{-1} A_{\mathfrak{s}_0}^\nu |\Delta_{12}\mathcal{R}_\nu|_s \,. \tag{4.70}
$$

Specializing (4.70) for $s = \mathfrak{s}_0$ and using (3.68), (2.20), (4.19), (4.23), we deduce

$$
|\Delta_{12}\mathcal{R}_{\nu+1}|_{\mathfrak{s}_0} \leq C(\varepsilon N_{\nu-1} N_\nu^{-\beta} + N_\nu^{2\tau+1} N_{\nu-1}^{-2\alpha}\varepsilon^2\gamma^{-1})\|u_1 - u_2\|_{\mathfrak{s}_0+\sigma_2} \leq \varepsilon N_\nu^{-\alpha}\|u_1 - u_2\|_{\mathfrak{s}_0+\sigma_2}
$$

for N_0 large and $\varepsilon\gamma^{-1}$ small. Next by (4.70) with $s = \mathfrak{s}_0 + \beta$

$$
|\Delta_{12}\mathcal{R}_\nu|_{\mathfrak{s}_0+\beta} \quad \overset{(4.19),(4.23),(4.14)}{\leq_{\mathfrak{s}_0+\beta}} \quad A_{\mathfrak{s}_0+\beta}^\nu \|u_1 - u_2\|_{\mathfrak{s}_0+\sigma_2} + |\Delta_{12}\mathcal{R}_\nu|_{\mathfrak{s}_0+\beta}
$$
$$
\overset{(4.19)(4.23)}{\leq} \quad C(\mathfrak{s}_0+\beta)\varepsilon N_{\nu-1}\|u_1 - u_2\|_{\mathfrak{s}_0+\sigma_2} \leq \varepsilon N_\nu \|u_1 - u_2\|_{\mathfrak{s}_0+\sigma_2}
$$

for N_0 large enough. Finally note that (4.24) is nothing but (4.55).

PROOF OF $(\mathbf{S4})_{\nu+1}$. We have to prove that, if $C\varepsilon N_\nu^\tau \|u_1 - u_2\|_{\mathfrak{s}_0+\sigma_2} \leq \rho$, then

$$
\lambda \in \Lambda_{\nu+1}^\gamma(u_1) \quad \Longrightarrow \quad \lambda \in \Lambda_{\nu+1}^{\gamma-\rho}(u_2)\,.
$$

Let $\lambda \in \Lambda_{\nu+1}^\gamma(u_1)$. Definition (4.17) and $(\mathbf{S4})_\nu$ (see (4.26)) imply that $\Lambda_{\nu+1}^\gamma(u_1) \subseteq \Lambda_\nu^\gamma(u_1) \subseteq \Lambda_\nu^{\gamma-\rho}(u_2)$. Hence $\lambda \in \Lambda_\nu^{\gamma-\rho}(u_2) \subset \Lambda_\nu^{\gamma/2}(u_2)$. Then, by $(\mathbf{S1})_\nu$, the eigenvalues $\mu_j^\nu(\lambda, u_2(\lambda))$ are well defined. Now (4.16) and the estimates (3.64), (4.25) (which holds because $\lambda \in \Lambda_\nu^\gamma(u_1) \cap \Lambda_\nu^{\gamma/2}(u_2)$) imply that

$$
|(\mu_j^\nu - \mu_k^\nu)(\lambda, u_2(\lambda)) - (\mu_j^\nu - \mu_k^\nu)(\lambda, u_1(\lambda))| \quad \leq \quad |(\mu_j^0 - \mu_k^0)(\lambda, u_2(\lambda)) - (\mu_j^0 - \mu_k^0)(\lambda, u_1(\lambda))|
$$
$$
+ 2\sup_{j\in\mathbb{Z}}|r_j^\nu(\lambda, u_2(\lambda)) - r_j^\nu(\lambda, u_1(\lambda))|
$$
$$
\leq \quad \varepsilon C|j^3 - k^3|\|u_2 - u_1\|_{\mathfrak{s}_0+\sigma_2}^{\sup}\,. \tag{4.71}
$$

Then we conclude that for all $|l| \leq N_\nu$, $j \neq k$, using the definition of $\Lambda_{\nu+1}^\gamma(u_1)$ (which is (4.17) with $\nu+1$ instead of ν) and (4.71),

$$
|\mathrm{i}\omega \cdot l + \mu_j^\nu(u_2) - \mu_k^\nu(u_2)| \quad \geq \quad |\mathrm{i}\omega \cdot l + \mu_j^\nu(u_1) - \mu_k^\nu(u_1)| - |(\mu_j^\nu - \mu_k^\nu)(u_2) - (\mu_j^\nu - \mu_k^\nu)(u_1)|
$$
$$
\geq \quad \gamma|j^3 - k^3|\langle l\rangle^{-\tau} - C\varepsilon|j^3 - k^3|\|u_1 - u_2\|_{\mathfrak{s}_0+\sigma_2}
$$
$$
\geq \quad (\gamma - \rho)|j^3 - k^3|\langle l\rangle^{-\tau}
$$

provided $C\varepsilon N_\nu^\tau \|u_1 - u_2\|_{\mathfrak{s}_0+\sigma_2} \leq \rho$. Hence $\lambda \in \Lambda_{\nu+1}^{\gamma-\rho}(u_2)$. This proves (4.26) at the step $\nu+1$.

4.2 Inversion of $\mathcal{L}(u)$

In (3.57) we have conjugated the linearized operator $\mathcal{L}$ to $\mathcal{L}_5$ defined in (3.55), namely $\mathcal{L} = \Phi_1 \mathcal{L}_5 \Phi_2^{-1}$. In Theorem 4.1 we have conjugated the operator $\mathcal{L}_5$ to the diagonal operator $\mathcal{L}_\infty$ in (4.7), namely $\mathcal{L}_5 = \Phi_\infty \mathcal{L}_\infty \Phi_\infty^{-1}$. As a consequence

$$
\mathcal{L} = W_1 \mathcal{L}_\infty W_2^{-1}, \quad W_i := \Phi_i \Phi_\infty, \quad i = 1, 2\,. \tag{4.72}
$$

We first prove that W_1, W_2 and their inverses are linear bijections of H^s. We take

$$
\gamma \leq \gamma_0/2, \quad \tau \geq \tau_0\,. \tag{4.73}
$$

Lemma 4.5. *Let $\mathfrak{s}_0 \leq s \leq q - \sigma - \beta - 3$ where β is defined in (4.1) and σ in (3.58). Let $u := u(\lambda)$ satisfy $\|u\|_{\mathfrak{s}_0+\sigma+\beta+3}^{\mathrm{Lip}(\gamma)} \leq 1$, and $\varepsilon\gamma^{-1} \leq \delta$ be small enough. Then W_i, $i = 1, 2$, satisfy, $\forall \lambda \in \Lambda_\infty^{2\gamma}(u)$,*

$$\|W_i h\|_s + \|W_i^{-1} h\|_s \leq C(s)\big(\|h\|_s + \|u\|_{s+\sigma+\beta}\|h\|_{\mathfrak{s}_0}\big), \tag{4.74}$$

$$\|W_i h\|_s^{\mathrm{Lip}(\gamma)} + \|W_i^{-1} h\|_s^{\mathrm{Lip}(\gamma)} \leq C(s)\big(\|h\|_{s+3}^{\mathrm{Lip}(\gamma)} + \|u\|_{s+\sigma+\beta+3}^{\mathrm{Lip}(\gamma)}\|h\|_{\mathfrak{s}_0+3}^{\mathrm{Lip}(\gamma)}\big). \tag{4.75}$$

In the reversible case (i.e. (1.13) holds), W_i, W_i^{-1}, $i = 1, 2$ are reversibility-preserving.

PROOF. The bound (4.74), resp. (4.75), follows by (4.8), (3.60), resp. (3.62), (2.12) and Lemma 6.5. In the reversible case $W_i^{\pm 1}$ are reversibility preserving because $\Phi_i^{\pm 1}$, $\Phi_\infty^{\pm 1}$ are reversibility preserving. ■

By (4.72) we are reduced to show that, $\forall \lambda \in \Lambda_\infty^{2\gamma}(u)$, the operator

$$\mathcal{L}_\infty := \mathrm{diag}_{j \in \mathbb{Z}}\{\mathrm{i}\lambda\bar{\omega} \cdot l + \mu_j^\infty(\lambda)\}, \quad \mu_j^\infty(\lambda) = -\mathrm{i}\big(m_3(\lambda)j^3 - m_1(\lambda)j\big) + r_j^\infty(\lambda)$$

is invertible, assuming (1.8) or the reversibility condition (1.13).

We introduce the following notation:

$$\Pi_C u := \frac{1}{(2\pi)^{\nu+1}} \int_{\mathbb{T}^{\nu+1}} u(\varphi, x)\,d\varphi dx, \quad \mathbb{P}u := u - \Pi_C u, \quad H_{00}^s := \{u \in H^s(\mathbb{T}^{\nu+1}) : \Pi_C u = 0\}. \tag{4.76}$$

If (1.8) holds, then the linearized operator $\mathcal{L}$ in (3.1) satisfies

$$\mathcal{L} : H^{s+3} \to H_{00}^s \tag{4.77}$$

(for $\mathfrak{s}_0 \leq s \leq q - 1$). In the reversible case (1.13)

$$\mathcal{L} : X \cap H^{s+3} \to Y \cap H^s \subset H_{00}^s. \tag{4.78}$$

Lemma 4.6. *Assume either (1.8) or the reversibility condition (1.13). Then the eigenvalue*

$$\mu_0^\infty(\lambda) = r_0^\infty(\lambda) = 0, \quad \forall \lambda \in \Lambda_\infty^{2\gamma}(u). \tag{4.79}$$

PROOF. Assume (1.8). If $r_0^\infty \neq 0$ then there exists a solution of $\mathcal{L}_\infty w = 1$, which is $w = 1/r_0^\infty$. Therefore, by (4.72),

$$\mathcal{L} W_2[1/r_0^\infty] = \mathcal{L} W_2 w = W_1 \mathcal{L}_\infty w = W_1[1]$$

which is a contradiction because $\Pi_C W_1[1] \neq 0$, for $\varepsilon\gamma^{-1}$ small enough, but the average $\Pi_C \mathcal{L} W_2[1/r_0^\infty] = 0$ by (4.77). In the reversible case $r_0^\infty = 0$ was proved in remark 4.3. ■

As a consequence of (4.79), the definition of $\Lambda_\infty^{2\gamma}$ in (4.6) (just specializing (4.6) with $k = 0$), and (1.2) (with γ and τ as in (4.73)), we deduce also the *first* order Melnikov non-resonance conditions

$$\forall \lambda \in \Lambda_\infty^{2\gamma}, \qquad \big|\mathrm{i}\lambda\bar{\omega} \cdot l + \mu_j^\infty(\lambda)\big| \geq 2\gamma \frac{\langle j \rangle^3}{\langle l \rangle^\tau}, \quad \forall (l, j) \neq (0, 0). \tag{4.80}$$

Lemma 4.7. (Invertibility of $\mathcal{L}_\infty$) *For all $\lambda \in \Lambda_\infty^{2\gamma}(u)$, for all $g \in H_{00}^s$ the equation $\mathcal{L}_\infty w = g$ has the unique solution with zero average*

$$\mathcal{L}_\infty^{-1} g(\varphi, x) := \sum_{(l,j)\neq(0,0)} \frac{g_{lj}}{\mathrm{i}\lambda\bar{\omega} \cdot l + \mu_j^\infty(\lambda)} e^{\mathrm{i}(l\cdot\varphi + jx)}. \tag{4.81}$$

For all Lipschitz family $g := g(\lambda) \in H_{00}^s$ we have

$$\big\|\mathcal{L}_\infty^{-1} g\big\|_s^{\mathrm{Lip}(\gamma)} \leq C\gamma^{-1} \|g\|_{s+2\tau+1}^{\mathrm{Lip}(\gamma)}. \tag{4.82}$$

In the reversible case, if $g \in Y$ then $\mathcal{L}_\infty^{-1} g \in X$.

PROOF. For all $\lambda \in \Lambda_\infty^{2\gamma}(u)$, by (4.80), formula (4.81) is well defined and

$$\left\| \mathcal{L}_\infty^{-1}(\lambda) g(\lambda) \right\|_s \ll \gamma^{-1} \left\| g(\lambda) \right\|_{s+\tau} . \tag{4.83}$$

Now we prove the Lipschitz estimate. For $\lambda_1, \lambda_2 \in \Lambda_\infty^{2\gamma}(u)$

$$\mathcal{L}_\infty^{-1}(\lambda_1) g(\lambda_1) - \mathcal{L}_\infty^{-1}(\lambda_2) g(\lambda_2) = \mathcal{L}_\infty^{-1}(\lambda_1)[g(\lambda_1) - g(\lambda_2)] + \left(\mathcal{L}_\infty^{-1}(\lambda_1) - \mathcal{L}_\infty^{-1}(\lambda_2) \right) g(\lambda_2) . \tag{4.84}$$

By (4.83)

$$\gamma \| \mathcal{L}_\infty^{-1}(\lambda_1)[g(\lambda_1) - g(\lambda_2)] \|_s \ll \| g(\lambda_1) - g(\lambda_2) \|_{s+\tau} \leq \gamma^{-1} \| g \|_{s+\tau}^{\mathrm{Lip}(\gamma)} |\lambda_1 - \lambda_2| . \tag{4.85}$$

Now we estimate the second term of (4.84). We simplify notations writing $g := g(\lambda_2)$ and $\delta_{lj} := \mathrm{i} \lambda \bar\omega \cdot l + \mu_j^\infty$.

$$\left(\mathcal{L}_\infty^{-1}(\lambda_1) - \mathcal{L}_\infty^{-1}(\lambda_2) \right) g = \sum_{(l,j) \neq (0,0)} \frac{\delta_{lj}(\lambda_2) - \delta_{lj}(\lambda_1)}{\delta_{lj}(\lambda_1) \delta_{lj}(\lambda_2)} g_{lj} e^{\mathrm{i}(l \cdot \varphi + jx)} . \tag{4.86}$$

The bound (4.5) imply $|\mu_j^\infty|^{\mathrm{lip}} \ll \varepsilon \gamma^{-1} |j|^3 \ll |j|^3$ and, using also (4.80),

$$\gamma \frac{|\delta_{lj}(\lambda_2) - \delta_{lj}(\lambda_1)|}{|\delta_{lj}(\lambda_1)||\delta_{lj}(\lambda_2)|} \ll \frac{(|l| + |j|^3) \langle l \rangle^{2\tau}}{\gamma \langle j \rangle^6} |\lambda_2 - \lambda_1| \ll \langle l \rangle^{2\tau+1} \gamma^{-1} |\lambda_2 - \lambda_1| . \tag{4.87}$$

Then (4.86) and (4.87) imply $\gamma \|(\mathcal{L}_\infty^{-1}(\lambda_2) - \mathcal{L}_\infty^{-1}(\lambda_1)) g\|_s \ll \gamma^{-1} \| g \|_{s+2\tau+1}^{\mathrm{Lip}(\gamma)} |\lambda_2 - \lambda_1|$ that, finally, with (4.83), (4.85), prove (4.82). The last statement follows by the property (4.37). ∎

In order to solve the equation $\mathcal{L} h = f$ we first prove the following lemma.

Lemma 4.8. *Let* $\mathfrak{s}_0 + \tau + 3 \leq s \leq q - \sigma - \beta - 3$. *Under the assumption* (1.8) *we have*

$$W_1(H_{00}^s) = H_{00}^s , \qquad W_1^{-1}(H_{00}^s) = H_{00}^s . \tag{4.88}$$

PROOF. It is sufficient to prove that $W_1(H_{00}^s) = H_{00}^s$ because the second equality of (4.88) follows applying the isomorphism W_1^{-1}. Let us give the proof of the inclusion

$$W_1(H_{00}^s) \subseteq H_{00}^s \tag{4.89}$$

(which is essentially algebraic). For any $g \in H_{00}^s$, let $w(\varphi, x) := \mathcal{L}_\infty^{-1} g \in H_{00}^{s-\tau}$ defined in (4.81). Then $h := W_2 w \in H^{s-\tau}$ satisfies

$$\mathcal{L} h \overset{(4.72)}{=} W_1 \mathcal{L}_\infty W_2^{-1} h = W_1 \mathcal{L}_\infty w = W_1 g .$$

By (4.77) we deduce that $W_1 g = \mathcal{L} h \in H_{00}^{s-\tau-3}$. Since $W_1 g \in H^s$ by Lemma 4.5, we conclude $W_1 g \in H^s \cap H_{00}^{s-\tau-3} = H_{00}^s$. The proof of (4.89) is complete.

It remains to prove that $H_{00}^s \setminus W_1(H_{00}^s) = \emptyset$. By contradiction, let $f \in H_{00}^s \setminus W_1(H_{00}^s)$. Let $g := W_1^{-1} f \in H^s$ by Lemma 4.5. Since $W_1 g = f \notin W_1(H_{00}^s)$, it follows that $g \notin H_{00}^s$ (otherwise it contradicts (4.89)), namely $c := \Pi_C g \neq 0$. Decomposing $g = c + \mathbb{P} g$ (recall (4.76)) and applying W_1, we get $W_1 g = c W_1[1] + W_1 \mathbb{P} g$. Hence

$$W_1[1] = c^{-1}(W_1 g - W_1 \mathbb{P} g) \in H_{00}^s$$

because $W_1 g = f \in H_{00}^s$ and $W_1 \mathbb{P} g \in W_1(H_{00}^s) \subseteq H_{00}^s$ by (4.89). However, $\Pi_C W_1[1] \neq 0$, a contradiction. ∎

Remark 4.7. *In the Hamiltonian case (which always satisfies* (1.8)*), the* $W_i(\varphi)$ *are maps of (a subspace of)* H_0^1 *so that Lemma 4.8 is automatic, and there is no need of Lemma 4.6.*

We may now prove the main result of sections 3 and 4.

Theorem 4.3. (Right inverse of $\mathcal{L}$) *Let*

$$\tau_1 := 2\tau + 7, \quad \mu := 4\tau + \sigma + \beta + 14, \tag{4.90}$$

where σ, β are defined in (3.58), (4.1) respectively. Let $u(\lambda)$, $\lambda \in \Lambda_o \subseteq \Lambda$, be a Lipschitz family with

$$\|u\|_{\mathfrak{s}_0+\mu}^{\mathrm{Lip}(\gamma)} \le 1. \tag{4.91}$$

Then there exists δ (depending on the data of the problem) such that if

$$\varepsilon\gamma^{-1} \le \delta,$$

and condition (1.8), resp. the reversibility condition (1.13), holds, then for all $\lambda \in \Lambda_\infty^{2\gamma}(u)$ defined in (4.6), the linearized operator $\mathcal{L} := \mathcal{L}(\lambda, u(\lambda))$ (see (3.1)) admits a right inverse on H_{00}^s, resp. $Y \cap H^s$. More precisely, for $\mathfrak{s}_0 \le s \le q - \mu$, for all Lipschitz family $f(\lambda) \in H_{00}^s$, resp. $Y \cap H^s$, the function

$$h := \mathcal{L}^{-1} f := W_2 \mathcal{L}_\infty^{-1} W_1^{-1} f \tag{4.92}$$

is a solution of $\mathcal{L}h = f$. In the reversible case, $\mathcal{L}^{-1} f \in X$. Moreover

$$\|\mathcal{L}^{-1} f\|_s^{\mathrm{Lip}(\gamma)} \le C(s)\gamma^{-1}\left(\|f\|_{s+\tau_1}^{\mathrm{Lip}(\gamma)} + \|u\|_{s+\mu}^{\mathrm{Lip}(\gamma)}\|f\|_{\mathfrak{s}_0}^{\mathrm{Lip}(\gamma)}\right). \tag{4.93}$$

PROOF. Given $f \in H_{00}^s$, resp. $f \in Y \cap H^s$, with s like in Lemma 4.8, the equation $\mathcal{L}h = f$ can be solved for h because $\Pi_C f = 0$. Indeed, by (4.72), the equation $\mathcal{L}h = f$ is equivalent to $\mathcal{L}_\infty W_2^{-1} h = W_1^{-1} f$ where $W_1^{-1} f \in H_{00}^s$ by Lemma 4.8, resp. $W_1^{-1} f \in Y \cap H^s$ being W_1^{-1} reversibility-preserving (Lemma 4.5). As a consequence, by Lemma 4.7, all the solutions of $\mathcal{L}h = f$ are

$$h = cW_2[1] + W_2\mathcal{L}_\infty^{-1} W_1^{-1} f, \quad c \in \mathbb{R}. \tag{4.94}$$

The solution (4.92) is the one with $c = 0$. In the reversible case, the fact that $\mathcal{L}^{-1} f \in X$ follows by (4.92) and the fact that W_i, W_i^{-1} are reversibility-preserving and $\mathcal{L}_\infty^{-1} : Y \to X$, see Lemma 4.7.

Finally (4.75), (4.82), (4.91) imply

$$\|\mathcal{L}^{-1} f\|_s^{\mathrm{Lip}(\gamma)} \le C(s)\gamma^{-1}\left(\|f\|_{s+2\tau+7}^{\mathrm{Lip}(\gamma)} + \|u\|_{s+2\tau+\sigma+\beta+7}^{\mathrm{Lip}(\gamma)}\|f\|_{\mathfrak{s}_0+2\tau+7}^{\mathrm{Lip}(\gamma)}\right)$$

and (4.93) follows using (6.2) with $b_0 = \mathfrak{s}_0$, $a_0 := \mathfrak{s}_0 + 2\tau + \sigma + \beta + 7$, $q = 2\tau + 7$, $p = s - \mathfrak{s}_0$. ∎

In the next section we apply Theorem 4.3 to deduce tame estimates for the inverse linearized operators at any step of the Nash-Moser scheme. The approximate solutions along the iteration will satisfy (4.91).

5 The Nash-Moser iteration

We define the finite-dimensional subspaces of trigonometric polynomials

$$H_n := \left\{ u \in L^2(\mathbb{T}^{\nu+1}) : u(\varphi, x) = \sum_{|(l,j)| \le N_n} u_{lj} e^{\mathrm{i}(l\cdot\varphi + jx)} \right\}$$

where $N_n := N_0^{\chi^n}$ (see (4.12)) and the corresponding orthogonal projectors

$$\Pi_n := \Pi_{N_n} : L^2(\mathbb{T}^{\nu+1}) \to H_n, \quad \Pi_n^\perp := I - \Pi_n.$$

The following smoothing properties hold: for all $\alpha, s \ge 0$,

$$\|\Pi_n u\|_{s+\alpha}^{\mathrm{Lip}(\gamma)} \le N_n^\alpha \|u\|_s^{\mathrm{Lip}(\gamma)}, \quad \forall u(\lambda) \in H^s; \quad \|\Pi_n^\perp u\|_s^{\mathrm{Lip}(\gamma)} \le N_n^{-\alpha} \|u\|_{s+\alpha}^{\mathrm{Lip}(\gamma)}, \quad \forall u(\lambda) \in H^{s+\alpha}, \tag{5.1}$$

where the function $u(\lambda)$ depends on the parameter λ in a Lipschitz way. The bounds (5.1) are the classical smoothing estimates for truncated Fourier series, which also hold with the norm $\|\cdot\|_s^{\mathrm{Lip}(\gamma)}$ defined in (2.2).

Let

$$F(u) := F(\lambda, u) := \lambda\bar{\omega} \cdot \partial_\varphi u + u_{xxx} + \varepsilon f(\varphi, x, u, u_x, u_{xx}, u_{xxx}). \qquad (5.2)$$

We define the constants

$$\kappa := 28 + 6\mu, \qquad \beta_1 := 50 + 11\mu, \qquad (5.3)$$

where μ is the loss of regularity in (4.90).

Theorem 5.1. (Nash-Moser) *Assume that $f \in C^q$, $q \geq \mathfrak{s}_0 + \mu + \beta_1$, satisfies the assumptions of Theorem 1.1 or Theorem 1.3. Let $0 < \gamma \leq \min\{\gamma_0, 1/48\}$, $\tau > \nu + 1$. Then there exist $\delta > 0$, $C_* > 0$, $N_0 \in \mathbb{N}$ (that may depend also on τ) such that, if $\varepsilon\gamma^{-1} < \delta$, then, for all $n \geq 0$:*

$(\mathcal{P}1)_n$ *there exists a function $u_n : \mathcal{G}_n \subseteq \Lambda \to H_n$, $\lambda \mapsto u_n(\lambda)$, with $\|u_n\|_{\mathfrak{s}_0+\mu}^{\mathrm{Lip}(\gamma)} \leq 1$, $u_0 := 0$, where $\mathcal{G}_n$ are Cantor like subsets of $\Lambda := [1/2, 3/2]$ defined inductively by: $\mathcal{G}_0 := \Lambda$,*

$$\mathcal{G}_{n+1} := \left\{ \lambda \in \mathcal{G}_n : |\mathrm{i}\omega \cdot l + \mu_j^\infty(u_n) - \mu_k^\infty(u_n)| \geq \frac{2\gamma_n|j^3 - k^3|}{\langle l \rangle^\tau}, \ \forall j, k \in \mathbb{Z}, \ l \in \mathbb{Z}^\nu \right\} \qquad (5.4)$$

where $\gamma_n := \gamma(1 + 2^{-n})$. In the reversible case, namely (1.13) holds, then $u_n(\lambda) \in X$.

The difference $h_n := u_n - u_{n-1}$, where, for convenience, $h_0 := 0$, satisfy

$$\|h_n\|_{\mathfrak{s}_0+\mu}^{\mathrm{Lip}(\gamma)} \leq C_*\varepsilon\gamma^{-1}N_n^{-\sigma_1}, \quad \sigma_1 := 18 + 2\mu. \qquad (5.5)$$

$(\mathcal{P}2)_n$ $\|F(u_n)\|_{\mathfrak{s}_0}^{\mathrm{Lip}(\gamma)} \leq C_*\varepsilon N_n^{-\kappa}$.

$(\mathcal{P}3)_n$ *(High norms).* $\|u_n\|_{\mathfrak{s}_0+\beta_1}^{\mathrm{Lip}(\gamma)} \leq C_*\varepsilon\gamma^{-1}N_n^\kappa$ *and* $\|F(u_n)\|_{\mathfrak{s}_0+\beta_1}^{\mathrm{Lip}(\gamma)} \leq C_*\varepsilon N_n^\kappa$.

$(\mathcal{P}4)_n$ *(Measure). The measure of the Cantor like sets satisfy*

$$|\mathcal{G}_0 \setminus \mathcal{G}_1| \leq C_*\gamma, \quad |\mathcal{G}_n \setminus \mathcal{G}_{n+1}| \leq \gamma C_* N_n^{-1}, \ n \geq 1. \qquad (5.6)$$

All the Lip norms are defined on $\mathcal{G}_n$.

PROOF. The proof of Theorem 5.1 is split into several steps. For simplicity, we denote $\| \ \|^{\mathrm{Lip}}$ by $\| \ \|$.

STEP 1: *prove* $(\mathcal{P}1,2,3)_0$. $(\mathcal{P}1)_0$ and the first inequality of $(\mathcal{P}3)_0$ are trivial because $u_0 = h_0 = 0$. $(\mathcal{P}2)_0$ and the second inequality of $(\mathcal{P}3)_0$ follow with $C_* \geq \max\{\|f(0)\|_{\mathfrak{s}_0}N_0^\kappa, \|f(0)\|_{\mathfrak{s}_0+\beta_1}N_0^{-\kappa}\}$.

STEP 2: *assume that $(\mathcal{P}1,2,3)_n$ hold for some $n \geq 0$, and prove $(\mathcal{P}1,2,3)_{n+1}$.* By $(\mathcal{P}1)_n$ we know that $\|u_n\|_{\mathfrak{s}_0+\mu} \leq 1$, namely condition (4.91) is satisfied. Hence, for $\varepsilon\gamma^{-1}$ small enough, Theorem 4.3 applies. Then, for all $\lambda \in \mathcal{G}_{n+1}$ defined in (5.4), the linearized operator

$$\mathcal{L}_n(\lambda) := \mathcal{L}(\lambda, u_n(\lambda)) = F'(\lambda, u_n(\lambda))$$

(see (3.1)) admits a right inverse for all $h \in H_{00}^s$, if condition (1.8) holds, respectively for $h \in Y \cap H^s$ if the reversibility condition (1.13) holds. Moreover (4.93) gives the estimates

$$\|\mathcal{L}_n^{-1}h\|_s \leq_s \gamma^{-1}\left(\|h\|_{s+\tau_1} + \|u_n\|_{s+\mu}\|h\|_{\mathfrak{s}_0}\right), \quad \forall h(\lambda), \qquad (5.7)$$

$$\|\mathcal{L}_n^{-1}h\|_{\mathfrak{s}_0} \leq \gamma^{-1}N_{n+1}^{\tau_1}\|h\|_{\mathfrak{s}_0}, \quad \forall h(\lambda) \in H_{n+1}, \qquad (5.8)$$

(use (5.1) and $\|u_n\|_{\mathfrak{s}_0+\mu} \leq 1$), for all Lipschitz map $h(\lambda)$. Then, for all $\lambda \in \mathcal{G}_{n+1}$, we define

$$u_{n+1} := u_n + h_{n+1} \in H_{n+1}, \quad h_{n+1} := -\Pi_{n+1}\mathcal{L}_n^{-1}\Pi_{n+1}F(u_n), \qquad (5.9)$$

which is well defined because, if condition (1.8) holds then $\Pi_{n+1}F(u_n) \in H_{00}^s$, and, respectively, if (1.13) holds, then $\Pi_{n+1}F(u_n) \in Y \cap H^s$ (hence in both cases $\mathcal{L}_n^{-1}\Pi_{n+1}F(u_n)$ exists). Note also that in the reversible case $h_{n+1} \in X$ and so $u_{n+1} \in X$.

Recalling (5.2) and that $\mathcal{L}_n := F'(u_n)$, we write

$$F(u_{n+1}) = F(u_n) + \mathcal{L}_n h_{n+1} + \varepsilon Q(u_n, h_{n+1}) \tag{5.10}$$

where

$$Q(u_n, h_{n+1}) := \mathcal{N}(u_n + h_{n+1}) - \mathcal{N}(u_n) - \mathcal{N}'(u_n) h_{n+1}, \quad \mathcal{N}(u) := f(\varphi, x, u, u_x, u_{xx}, u_{xxx}).$$

With this definition,

$$F(u) = L_\omega u + \varepsilon \mathcal{N}(u), \quad F'(u)h = L_\omega h + \varepsilon \mathcal{N}'(u)h, \quad L_\omega := \omega \cdot \partial_\varphi + \partial_{xxx}.$$

By (5.10) and (5.9) we have

$$
\begin{aligned}
F(u_{n+1}) &= F(u_n) - \mathcal{L}_n \Pi_{n+1} \mathcal{L}_n^{-1} \Pi_{n+1} F(u_n) + \varepsilon Q(u_n, h_{n+1}) \\
&= \Pi_{n+1}^\perp F(u_n) + \mathcal{L}_n \Pi_{n+1}^\perp \mathcal{L}_n^{-1} \Pi_{n+1} F(u_n) + \varepsilon Q(u_n, h_{n+1}) \\
&= \Pi_{n+1}^\perp F(u_n) + \Pi_{n+1}^\perp \mathcal{L}_n \mathcal{L}_n^{-1} \Pi_{n+1} F(u_n) + [\mathcal{L}_n, \Pi_{n+1}^\perp] \mathcal{L}_n^{-1} \Pi_{n+1} F(u_n) + \varepsilon Q(u_n, h_{n+1}) \\
&= \Pi_{n+1}^\perp F(u_n) + \varepsilon [\mathcal{N}'(u_n), \Pi_{n+1}^\perp] \mathcal{L}_n^{-1} \Pi_{n+1} F(u_n) + \varepsilon Q(u_n, h_{n+1})
\end{aligned}
\tag{5.11}
$$

where we have gained an extra ε from the commutator

$$[\mathcal{L}_n, \Pi_{n+1}^\perp] = [L_\omega + \varepsilon \mathcal{N}'(u_n), \Pi_{n+1}^\perp] = \varepsilon [\mathcal{N}'(u_n), \Pi_{n+1}^\perp].$$

Lemma 5.1. *Set*

$$U_n := \|u_n\|_{\mathfrak{s}_0 + \beta_1} + \gamma^{-1} \|F(u_n)\|_{\mathfrak{s}_0 + \beta_1}, \qquad w_n := \gamma^{-1} \|F(u_n)\|_{\mathfrak{s}_0}. \tag{5.12}$$

There exists $C_0 := C(\tau_1, \mu, \nu, \beta_1) > 0$ such that

$$w_{n+1} \leq C_0 N_{n+1}^{-\beta_1 + \mu'} U_n (1 + w_n) + C_0 N_{n+1}^{6+2\mu} w_n^2, \qquad U_{n+1} \leq C_0 N_{n+1}^{9+2\mu} (1 + w_n)^2 U_n. \tag{5.13}$$

PROOF. The operators $\mathcal{N}'(u_n)$ and $Q(u_n, \cdot)$ satisfy the following tame estimates:

$$\|Q(u_n, h)\|_s \lesssim_s \|h\|_{\mathfrak{s}_0 + 3} \big(\|h\|_{s+3} + \|u_n\|_{s+3} \|h\|_{\mathfrak{s}_0 + 3} \big) \quad \forall h(\lambda), \tag{5.14}$$

$$\|Q(u_n, h)\|_{\mathfrak{s}_0} \leq N_{n+1}^6 \|h\|_{\mathfrak{s}_0}^2 \quad \forall h(\lambda) \in H_{n+1}, \tag{5.15}$$

$$\|\mathcal{N}'(u_n)h\|_s \lesssim_s \|h\|_{s+3} + \|u_n\|_{s+3} \|h\|_{\mathfrak{s}_0 + 3} \quad \forall h(\lambda), \tag{5.16}$$

where $h(\lambda)$ depends on the parameter λ in a Lipschitz way. The bounds (5.14) and (5.16) follow by 6.2(i) and Lemma 6.3. (5.15) is simply (5.14) at $s = \mathfrak{s}_0$, using that $\|u_n\|_{\mathfrak{s}_0 + 3} \leq 1$, $u_n, h_{n+1} \in H_{n+1}$ and the smoothing (5.1).

By (5.7) and (5.16), the term (in (5.11)) $R_n := [\mathcal{N}'(u_n), \Pi_{n+1}^\perp] \mathcal{L}_n^{-1} \Pi_{n+1} F(u_n)$ satisfies, using also that $u_n \in H_n$ and (5.1),

$$\|R_n\|_s \lesssim_s \gamma^{-1} N_{n+1}^{\mu'} \big(\|F(u_n)\|_s + \|u_n\|_s \|F(u_n)\|_{\mathfrak{s}_0} \big), \quad \mu' := 3 + \mu, \tag{5.17}$$

$$\|R_n\|_{\mathfrak{s}_0 + \beta_1} \lesssim_{\mathfrak{s}_0 + \beta_1} \gamma^{-1} N_{n+1}^{-\beta_1 + \mu'} \big(\|F(u_n)\|_{\mathfrak{s}_0 + \beta_1} + \|u_n\|_{\mathfrak{s}_0 + \beta_1} \|F(u_n)\|_{\mathfrak{s}_0} \big), \tag{5.18}$$

because $\mu \geq \tau_1 + 3$. In proving (5.17) and (5.18), we have simply estimated $\mathcal{N}'(u_n) \Pi_{n+1}^\perp$ and $\Pi_{n+1}^\perp \mathcal{N}'(u_n)$ separately, without using the commutator structure.

From the definition (5.9) of h_{n+1}, using (5.7), (5.8) and (5.1), we get

$$\|h_{n+1}\|_{\mathfrak{s}_0 + \beta_1} \lesssim_{\mathfrak{s}_0 + \beta_1} \gamma^{-1} N_{n+1}^\mu \big(\|F(u_n)\|_{\mathfrak{s}_0 + \beta_1} + \|u_n\|_{\mathfrak{s}_0 + \beta_1} \|F(u_n)\|_{\mathfrak{s}_0} \big), \tag{5.19}$$

$$\|h_{n+1}\|_{\mathfrak{s}_0} \lesssim_{\mathfrak{s}_0} \gamma^{-1} N_{n+1}^\mu \|F(u_n)\|_{\mathfrak{s}_0} \tag{5.20}$$

35

because $\mu \geq \tau_1$. Then

$$\|u_{n+1}\|_{\mathfrak{s}_0+\beta_1} \overset{(5.9)}{\leq} \|u_n\|_{\mathfrak{s}_0+\beta_1} + \|h_{n+1}\|_{\mathfrak{s}_0+\beta_1}$$

$$\overset{(5.19)}{\leq}_{\mathfrak{s}_0+\beta_1} \|u_n\|_{\mathfrak{s}_0+\beta_1} \left(1 + \gamma^{-1} N_{n+1}^\mu \|F(u_n)\|_{\mathfrak{s}_0}\right) + \gamma^{-1} N_{n+1}^\mu \|F(u_n)\|_{\mathfrak{s}_0+\beta_1}. \qquad (5.21)$$

Formula (5.11) for $F(u_{n+1})$, and (5.18), (5.15), (5.20), $\varepsilon\gamma^{-1} \leq 1$, (5.1), imply

$$\|F(u_{n+1})\|_{\mathfrak{s}_0} \leq_{\mathfrak{s}_0+\beta_1} N_{n+1}^{-\beta_1+\mu'} \left(\|F(u_n)\|_{\mathfrak{s}_0+\beta_1} + \|u_n\|_{\mathfrak{s}_0+\beta_1} \|F(u_n)\|_{\mathfrak{s}_0}\right) + \varepsilon\gamma^{-2} N_{n+1}^{6+2\mu} \|F(u_n)\|_{\mathfrak{s}_0}^2. \qquad (5.22)$$

Similarly, using the "high norm" estimates (5.17), (5.14), (5.19), (5.20), $\varepsilon\gamma^{-1} \leq 1$ and (5.1),

$$\|F(u_{n+1})\|_{\mathfrak{s}_0+\beta_1} \leq_{\mathfrak{s}_0+\beta_1} \left(\|F(u_n)\|_{\mathfrak{s}_0+\beta_1} + \|u_n\|_{\mathfrak{s}_0+\beta_1} \|F(u_n)\|_{\mathfrak{s}_0}\right) \left(1 + N_{n+1}^{\mu'} + N_{n+1}^{9+2\mu} \gamma^{-1} \|F(u_n)\|_{\mathfrak{s}_0}\right). \qquad (5.23)$$

By (5.21), (5.22) and (5.23) we deduce (5.13). ∎

By $(\mathcal{P}2)_n$ we deduce, for $\varepsilon\gamma^{-1}$ small, that (recall the definition on w_n in (5.12))

$$w_n \leq \varepsilon\gamma^{-1} C_* N_n^{-\kappa} \leq 1, \qquad (5.24)$$

Then, by the second inequality in (5.13), (5.24), $(\mathcal{P}3)_n$ (recall the definition on U_n in (5.12)) and the choice of κ in (5.3), we deduce $U_{n+1} \leq C_*\varepsilon\gamma^{-1} N_{n+1}^\kappa$, for N_0 large enough. This proves $(\mathcal{P}3)_{n+1}$.

Next, by the first inequality in (5.13), (5.24), $(\mathcal{P}2)_n$ (recall the definition on w_n in (5.12)) and (5.3), we deduce $w_{n+1} \leq C_*\varepsilon\gamma^{-1} N_{n+1}^\kappa$, for N_0 large, $\varepsilon\gamma^{-1}$ small. This proves $(\mathcal{P}2)_{n+1}$.

The bound (5.5) at the step $n+1$ follows by (5.20) and $(\mathcal{P}2)_n$ (and (5.3)). Then

$$\|u_{n+1}\|_{\mathfrak{s}_0+\mu} \leq \|u_0\|_{\mathfrak{s}_0+\mu} + \sum_{k=1}^{n+1} \|h_k\|_{\mathfrak{s}_0+\mu} \leq \sum_{k=1}^{\infty} C_*\varepsilon\gamma^{-1} N_k^{-\sigma_1} \leq 1$$

for $\varepsilon\gamma^{-1}$ small enough. As a consequence $(\mathcal{P}1,2,3)_{n+1}$ hold.

STEP 3: *prove* $(\mathcal{P}4)_n$, $n \geq 0$. For all $n \geq 0$,

$$\mathcal{G}_n \setminus \mathcal{G}_{n+1} = \bigcup_{l\in\mathbb{Z}^\nu, j,k\in\mathbb{Z}} R_{ljk}(u_n) \qquad (5.25)$$

where

$$R_{ljk}(u_n) := \left\{\lambda \in \mathcal{G}_n : \left|\mathrm{i}\lambda\bar{\omega}\cdot l + \mu_j^\infty(\lambda, u_n(\lambda)) - \mu_k^\infty(\lambda, u_n(\lambda))\right| < 2\gamma_n|j^3-k^3|\langle l\rangle^{-\tau}\right\}. \qquad (5.26)$$

Notice that, by the definition (5.26), $R_{ljk}(u_n) = \emptyset$ for $j = k$. Then we can suppose in the sequel that $j \neq k$. We divide the estimate into some lemmata.

Lemma 5.2. *For $\varepsilon\gamma^{-1}$ small enough, for all $n \geq 0$, $|l| \leq N_n$,*

$$R_{ljk}(u_n) \subseteq R_{ljk}(u_{n-1}). \qquad (5.27)$$

PROOF. We claim that, for all $j, k \in \mathbb{Z}$,

$$\left|(\mu_j^\infty - \mu_k^\infty)(u_n) - (\mu_j^\infty - \mu_k^\infty)(u_{n-1})\right| \leq C\varepsilon|j^3-k^3|N_n^{-\alpha}, \quad \forall\lambda \in \mathcal{G}_n, \qquad (5.28)$$

where $\mu_j^\infty(u_n) := \mu_j^\infty(\lambda, u_n(\lambda))$ and α is defined in (4.13). Before proving (5.28) we show how it implies (5.27). For all $j \neq k$, $|l| \leq N_n$, $\lambda \in \mathcal{G}_n$, by (5.28)

$$\left|\mathrm{i}\lambda\bar{\omega}\cdot l + \mu_j^\infty(u_n) - \mu_k^\infty(u_n)\right| \geq \left|\mathrm{i}\lambda\bar{\omega}\cdot l + \mu_j^\infty(u_{n-1}) - \mu_k^\infty(u_{n-1})\right| - \left|(\mu_j^\infty - \mu_k^\infty)(u_n) - (\mu_j^\infty - \mu_k^\infty)(u_{n-1})\right|$$

$$\geq 2\gamma_{n-1}|j^3-k^3|\langle l\rangle^{-\tau} - C\varepsilon|j^3-k^3|N_n^{-\alpha} \geq 2\gamma_n|j^3-k^3|\langle l\rangle^{-\tau}$$

36

for $C\varepsilon\gamma^{-1}N_n^{\tau-\alpha}\,2^{n+1} \le 1$ (recall that $\gamma_n := \gamma(1+2^{-n})$), which implies (5.27).

PROOF OF (5.28). By (4.4),

$$(\mu_j^\infty - \mu_k^\infty)(u_n) - (\mu_j^\infty - \mu_k^\infty)(u_{n-1}) = -\mathrm{i}\big[m_3(u_n) - m_3(u_{n-1})\big](j^3 - k^3) + \mathrm{i}\big[m_1(u_n) - m_1(u_{n-1})\big](j - k)$$
$$+ r_j^\infty(u_n) - r_j^\infty(u_{n-1}) - \big(r_k^\infty(u_n) - r_k^\infty(u_{n-1})\big) \tag{5.29}$$

where $m_3(u_n) := m_3(\lambda, u_n(\lambda))$ and similarly for m_1, r_j^∞. We first apply Theorem 4.2-$(\mathbf{S4})_\nu$ with $\nu = n+1$, $\gamma = \gamma_{n-1}$, $\gamma - \rho = \gamma_n$, and u_1, u_2, replaced, respectively, by u_{n-1}, u_n, in order to conclude that

$$\Lambda_{n+1}^{\gamma_{n-1}}(u_{n-1}) \subseteq \Lambda_{n+1}^{\gamma_n}(u_n). \tag{5.30}$$

The smallness condition in (4.26) is satisfied because $\sigma_2 < \mu$ (see definitions (4.13), (4.90)) and so

$$\varepsilon C N_n^\tau \|u_n - u_{n-1}\|_{\mathfrak{s}_0+\sigma_2} \le \varepsilon C N_n^\tau \|u_n - u_{n-1}\|_{\mathfrak{s}_0+\mu} \overset{(5.5)}{\le} \varepsilon^2 \gamma^{-1} C C_* N_n^{\tau-\sigma_1} \le \gamma_{n-1} - \gamma_n =: \rho = \gamma 2^{-n}$$

for $\varepsilon\gamma^{-1}$ small enough, because $\sigma_1 > \tau$ (see (5.5), (4.90)). Then, by the definitions (5.4) and (4.6), we have

$$\mathcal{G}_n := \mathcal{G}_{n-1} \cap \Lambda_\infty^{2\gamma_{n-1}}(u_{n-1}) \overset{(4.35)}{\subseteq} \bigcap_{\nu \ge 0} \Lambda_\nu^{\gamma_{n-1}}(u_{n-1}) \subset \Lambda_{n+1}^{\gamma_{n-1}}(u_{n-1}) \overset{(5.30)}{\subseteq} \Lambda_{n+1}^{\gamma_n}(u_n).$$

Next, for all $\lambda \in \mathcal{G}_n \subset \Lambda_{n+1}^{\gamma_{n-1}}(u_{n-1}) \cap \Lambda_{n+1}^{\gamma_n}(u_n)$ both $r_j^{n+1}(u_{n-1})$ and $r_j^{n+1}(u_n)$ are well defined, and we deduce by Theorem 4.2-$(\mathbf{S3})_\nu$ with $\nu = n+1$, that

$$|r_j^{n+1}(u_n) - r_j^{n+1}(u_{n-1})| \overset{(4.25)}{\prec} \varepsilon \|u_{n-1} - u_n\|_{\mathfrak{s}_0+\sigma_2}. \tag{5.31}$$

Moreover (4.34) (with $\nu = n+1$) and (3.66) imply that

$$|r_j^\infty(u_{n-1}) - r_j^{n+1}(u_{n-1})| + |r_j^\infty(u_n) - r_j^{n+1}(u_n)| \quad \prec \quad \varepsilon(1 + \|u_{n-1}\|_{\mathfrak{s}_0+\beta+\sigma} + \|u_n\|_{\mathfrak{s}_0+\beta+\sigma})N_n^{-\alpha}$$
$$\prec \quad \varepsilon N_n^{-\alpha} \tag{5.32}$$

because $\sigma + \beta < \mu$ and $\|u_{n-1}\|_{\mathfrak{s}_0+\mu} + \|u_n\|_{\mathfrak{s}_0+\mu} \le 2$ by $(\mathbf{S1})_{n-1}$ and $(\mathbf{S1})_n$. Therefore, for all $\lambda \in \mathcal{G}_n$, $\forall j \in \mathbb{Z}$,

$$\left|r_j^\infty(u_n) - r_j^\infty(u_{n-1})\right| \le \left|r_j^{n+1}(u_n) - r_j^{n+1}(u_{n-1})\right| + |r_j^\infty(u_n) - r_j^{n+1}(u_n)| + |r_j^\infty(u_{n-1}) - r_j^{n+1}(u_{n-1})|$$
$$\overset{(5.31),(5.32)}{\prec} \varepsilon \|u_n - u_{n-1}\|_{\mathfrak{s}_0+\sigma_2} + \varepsilon N_n^{-\alpha} \overset{(5.5)}{\prec} \varepsilon N_n^{-\alpha} \tag{5.33}$$

because $\sigma_1 > \alpha$ (see (4.13), (5.5)). Finally (5.29), (5.33), (3.64), $\|u_n\|_{\mathfrak{s}_0+\mu} \le 1$, imply (5.28). ∎

By definition, $R_{ljk}(u_n) \subset \mathcal{G}_n$ (see (5.26)) and, by (5.27), for all $|l| \le N_n$, we have $R_{ljk}(u_n) \subseteq R_{ljk}(u_{n-1})$. On the other hand $R_{ljk}(u_{n-1}) \cap \mathcal{G}_n = \emptyset$, see (5.4). As a consequence, $\forall |l| \le N_n$, $R_{ljk}(u_n) = \emptyset$, and

$$\mathcal{G}_n \setminus \mathcal{G}_{n+1} \overset{(5.25)}{\subseteq} \bigcup_{|l|>N_n, j, k \in \mathbb{Z}} R_{ljk}(u_n), \quad \forall n \ge 1. \tag{5.34}$$

Lemma 5.3. *Let $n \ge 0$. If $R_{ljk}(u_n) \ne \emptyset$, then $|j^3 - k^3| \le 8|\bar{\omega} \cdot l|$.*

PROOF. If $R_{ljk}(u_n) \ne \emptyset$ then there exists $\lambda \in \Lambda$ such that $|\mathrm{i}\lambda\bar{\omega} \cdot l + \mu_j^\infty(\lambda, u_n(\lambda)) - \mu_k^\infty(\lambda, u_n(\lambda))| < 2\gamma_n|j^3 - k^3|\langle l\rangle^{-\tau}$ and, therefore,

$$|\mu_j^\infty(\lambda, u_n(\lambda)) - \mu_k^\infty(\lambda, u_n(\lambda))| < 2\gamma_n|j^3 - k^3|\langle l\rangle^{-\tau} + 2|\bar{\omega} \cdot l|. \tag{5.35}$$

Moreover, by (4.4), (3.63), (4.5), for ε small enough,

$$|\mu_j^\infty - \mu_k^\infty| \ge |m_3||j^3 - k^3| - |m_1||j - k| - |r_j^\infty| - |r_k^\infty| \ge \frac{1}{2}|j^3 - k^3| - C\varepsilon|j - k| - C\varepsilon \ge \frac{1}{3}|j^3 - k^3| \tag{5.36}$$

if $j \ne k$. Since $\gamma_n \le 2\gamma$ for all $n \ge 0$, $\gamma \le 1/48$, by (5.35) and (5.36) we get

$$2|\bar{\omega} \cdot l| \ge \Big(\frac{1}{3} - \frac{4\gamma}{\langle l\rangle^\tau}\Big)|j^3 - k^3| \ge \frac{1}{4}|j^3 - k^3|$$

proving the Lemma. ∎

Lemma 5.4. *For all $n \geq 0$,*

$$|R_{ljk}(u_n)| \leq C\gamma \langle l \rangle^{-\tau}. \tag{5.37}$$

PROOF. Consider the function $\phi : \Lambda \to \mathbb{C}$ defined by

$$
\begin{aligned}
\phi(\lambda) &:= & \mathrm{i}\lambda\bar{\omega} \cdot l + \mu_j^\infty(\lambda) - \mu_k^\infty(\lambda) \\
&\overset{(4.4)}{=} & \mathrm{i}\lambda\bar{\omega} \cdot l - \mathrm{i}\tilde{m}_3(\lambda)(j^3 - k^3) + \mathrm{i}\tilde{m}_1(\lambda)(j - k) + r_j^\infty(\lambda) - r_k^\infty(\lambda)
\end{aligned}
$$

where $\tilde{m}_3(\lambda)$, $\tilde{m}_1(\lambda)$, $r_j^\infty(\lambda)$, $\mu_j^\infty(\lambda)$, are defined for all $\lambda \in \Lambda$ and satisfy (4.5) by $\|u_n\|_{\mathfrak{s}_0+\mu,\mathcal{G}_n}^{\mathrm{Lip}(\gamma)} \leq 1$ (see $(\mathcal{P}1)_n$). Recalling $|\cdot|^{\mathrm{lip}} \leq \gamma^{-1}|\cdot|^{\mathrm{Lip}(\gamma)}$ and using (4.5)

$$|\mu_j^\infty - \mu_k^\infty|^{\mathrm{lip}} \leq |\tilde{m}_3|^{\mathrm{lip}}|j^3 - k^3| + |\tilde{m}_1|^{\mathrm{lip}}|j - k| + |r_j^\infty|^{\mathrm{lip}} + |r_k^\infty|^{\mathrm{lip}} \leq C\varepsilon\gamma^{-1}|j^3 - k^3|. \tag{5.38}$$

Moreover Lemma 5.3 implies that, $\forall \lambda_1, \lambda_2 \in \Lambda$,

$$|\phi(\lambda_1) - \phi(\lambda_2)| \geq \left(|\bar{\omega} \cdot l| - |\mu_j^\infty - \mu_k^\infty|^{\mathrm{lip}}\right)|\lambda_1 - \lambda_2| \overset{(5.38)}{\geq} \left(\frac{1}{8} - C\varepsilon\gamma^{-1}\right)|j^3 - k^3||\lambda_1 - \lambda_2| \geq \frac{|j^3 - k^3|}{9}|\lambda_1 - \lambda_2|$$

for $\varepsilon\gamma^{-1}$ small enough. Hence

$$|R_{ljk}(u_n)| \leq \frac{4\gamma_n|j^3 - k^3|}{\langle l \rangle^\tau} \frac{9}{|j^3 - k^3|} \leq \frac{72\gamma}{\langle l \rangle^\tau},$$

which is (5.37). ∎

Now we prove $(\mathcal{P}4)_0$. We observe that, for each fixed l, all the indices j, k such that $R_{ljk}(0) \neq \emptyset$ are confined in the ball $j^2 + k^2 \leq 16|\bar{\omega}||l|$, because

$$|j^3 - k^3| = |j - k||j^2 + jk + k^2| \geq j^2 + k^2 - |jk| \geq \frac{1}{2}(j^2 + k^2), \quad \forall j, k \in \mathbb{Z}, \ j \neq k,$$

and $|j^3 - k^3| \leq 8|\bar{\omega}||l|$ by Lemma 5.3. As a consequence

$$|\mathcal{G}_0 \setminus \mathcal{G}_1| \overset{(5.25)}{=} \left| \bigcup_{l,j,k} R_{ljk}(0) \right| \leq \sum_{l \in \mathbb{Z}^\nu} \sum_{j^2 + k^2 \leq 16|\bar{\omega}||l|} |R_{ljk}(0)| \overset{(5.37)}{<} \sum_{l \in \mathbb{Z}^\nu} \gamma \langle l \rangle^{-\tau+1} = C\gamma$$

if $\tau > \nu + 1$. Thus the first estimate in (5.6) is proved, taking a larger C_* if necessary.

Finally, $(\mathcal{P}4)_n$ for $n \geq 1$, follows by

$$
\begin{aligned}
|\mathcal{G}_n \setminus \mathcal{G}_{n+1}| &\overset{(5.34)}{\leq} \sum_{|l| > N_n, |j|, |k| \leq C|l|^{1/2}} |R_{ljk}(u_n)| \overset{(5.37)}{<} \sum_{|l| > N_n, |j|, |k| \leq C|l|^{1/2}} \gamma \langle l \rangle^{-\tau} \\
&< \sum_{|l| > N_n} \gamma \langle l \rangle^{-\tau+1} < \gamma N_n^{-\tau+\nu} \leq C\gamma N_n^{-1}
\end{aligned}
$$

and (5.6) is proved. The proof of Theorem 5.1 is complete. ∎

5.1   Proof of Theorems 1.1, 1.2, 1.3, 1.4 and 1.5

PROOF OF THEOREMS 1.1, 1.2, 1.3. Assume that $f \in C^q$ satisfies the assumptions in Theorem 1.1 or in Theorem 1.3 with a smoothness exponent $q := q(\nu) \geq \mathfrak{s}_0 + \mu + \beta_1$ which depends only on ν once we have fixed $\tau := \nu + 2$ (recall that $\mathfrak{s}_0 := (\nu + 2)/2$, β_1 is defined in (5.3) and μ in (4.90)).

For $\gamma = \varepsilon^a$, $a \in (0, 1)$ the smallness condition $\varepsilon\gamma^{-1} = \varepsilon^{1-a} < \delta$ of Theorem 5.1 is satisfied. Hence on the Cantor set $\mathcal{G}_\infty := \cap_{n \geq 0}\mathcal{G}_n$, the sequence $u_n(\lambda)$ is well defined and converges in norm $\|\cdot\|_{\mathfrak{s}_0+\mu,\mathcal{G}_\infty}^{\mathrm{Lip}(\gamma)}$ (see (5.5)) to a solution $u_\infty(\lambda)$ of

$$F(\lambda, u_\infty(\lambda)) = 0 \quad \text{with} \quad \sup_{\lambda \in \mathcal{G}_\infty} \|u_\infty(\lambda)\|_{\mathfrak{s}_0+\mu} \leq C\varepsilon\gamma^{-1} = C\varepsilon^{1-a},$$

namely $u_\infty(\lambda)$ is a solution of the perturbed KdV equation (1.4) with $\omega = \lambda\bar\omega$. Moreover, by (5.6), the measure of the complementary set satisfies

$$|\Lambda \setminus \mathcal{G}_\infty| \leq \sum_{n \geq 0} |\mathcal{G}_n \setminus \mathcal{G}_{n+1}| \leq C\gamma + \sum_{n \geq 1} \gamma C N_n^{-1} \leq C\gamma = C\varepsilon^a \,,$$

proving (1.9). The proof of Theorem 1.1 is complete. In order to finish the proof of Theorems 1.2 or 1.3, it remains to prove the linear stability of the solution, namely Theorem 1.5.

PROOF OF THEOREM 1.4. Part (i) follows by (4.72), Lemma 4.5, Theorem 4.1 (applied to the solution $u_\infty(\lambda)$) with the exponents $\bar\sigma := \sigma + \beta + 3$, $\Lambda_\infty(u) := \Lambda_\infty^{2\gamma}(u)$, see (4.6). Part (ii) follows by the dynamical interpretation of the conjugation procedure, as explained in section 2.2. Explicitely, in sections 3 and 4, we have proved that

$$\mathcal{L} = \mathcal{A}B\rho W \mathcal{L}_\infty W^{-1} B^{-1} \mathcal{A}^{-1}, \quad W := \mathcal{MTS}\Phi_\infty \,.$$

By the arguments in Section 2.2 we deduce that a curve $h(t)$ in the phase space H_x^s is a solution of the dynamical system (1.19) if and only if the transformed curve

$$v(t) := W^{-1}(\omega t) B^{-1} \mathcal{A}^{-1}(\omega t) h(t) \tag{5.39}$$

(see notation (2.18), Lemma 3.3, (4.9)) is a solution of the constant coefficients dynamical system (1.20).

PROOF OF THEOREM 1.5. If all μ_j are purely imaginary, the Sobolev norm of the solution $v(t)$ of (1.20) is constant in time, see (1.21). We now show that also the Sobolev norm of the solution $h(t)$ in (5.39) does not grow in time. For each $t \in \mathbb{R}$, $\mathcal{A}(\omega t)$ and $W(\omega t)$ are transformations of the phase space H_x^s that depend quasi-periodically on time, and satisfy, by (3.69), (3.71), (4.9),

$$\|\mathcal{A}^{\pm 1}(\omega t) g\|_{H_x^s} + \|W^{\pm 1}(\omega t) g\|_{H_x^s} \leq C(s) \|g\|_{H_x^s}, \quad \forall t \in \mathbb{R}, \ \forall g = g(x) \in H_x^s, \tag{5.40}$$

where the constant $C(s)$ depends on $\|u\|_{s+\sigma+\beta+\mathfrak{s}_0} < +\infty$. Moreover, the transformation B is a quasi-periodic reparametrization of the time variable (see (2.25)), namely

$$Bf(t) = f(\psi(t)) = f(\tau), \quad B^{-1} f(\tau) = f(\psi^{-1}(\tau)) = f(t) \quad \forall f : \mathbb{R} \to H_x^s, \tag{5.41}$$

where $\tau = \psi(t) := t + \alpha(\omega t)$, $t = \psi^{-1}(\tau) = \tau + \tilde\alpha(\omega\tau)$ and α, $\tilde\alpha$ are defined in Section 3.2. Thus

$$\|h(t)\|_{H_x^s} \overset{(5.39)}{=} \|\mathcal{A}(\omega t) B W(\omega t) v(t)\|_{H_x^s} \overset{(5.40)}{\leq} C(s)\|B W(\omega t) v(t)\|_{H_x^s} \overset{(5.41)}{=} C(s)\|W(\omega\tau) v(\tau)\|_{H_x^s}$$

$$\overset{(5.40)}{\leq} C(s)\|v(\tau)\|_{H_x^s} \overset{(1.21)}{=} C(s)\|v(\tau_0)\|_{H_x^s} \overset{(5.39)}{=} C(s)\|W^{-1}(\omega\tau_0) B^{-1} \mathcal{A}^{-1}(\omega\tau_0) h(\tau_0)\|_{H_x^s}$$

$$\overset{(5.40)}{\leq} C(s)\|B^{-1}\mathcal{A}^{-1}(\omega\tau_0) h(\tau_0)\|_{H_x^s} \overset{(5.41)}{=} C(s)\|\mathcal{A}^{-1}(0) h(0)\|_{H_x^s} \overset{(5.40)}{\leq} C(s)\|h(0)\|_{H_x^s}$$

having chosen $\tau_0 := \psi(0) = \alpha(0)$ (in the reversible case, α is an odd function, and so $\alpha(0) = 0$). Hence (1.22) is proved. To prove (1.23), we collect the estimates (3.70), (3.72), (4.9) into

$$\|(\mathcal{A}^{\pm 1}(\omega t) - I) g\|_{H_x^s} + \|(W^{\pm 1}(\omega t) - I) g\|_{H_x^s} \leq \varepsilon\gamma^{-1} C(s) \|g\|_{H_x^{s+1}}, \quad \forall t \in \mathbb{R}, \ \forall g \in H_x^s, \tag{5.42}$$

where the constant $C(s)$ depends on $\|u\|_{s+\sigma+\beta+\mathfrak{s}_0}$. Thus

$$\|h(t)\|_{H_x^s} \overset{(5.39)}{=} \|\mathcal{A}(\omega t) B W(\omega t) v(t)\|_{H_x^s} \leq \|B W(\omega t) v(t)\|_{H_x^s} + \|(\mathcal{A}(\omega t) - I) B W(\omega t) v(t)\|_{H_x^s}$$

$$\overset{(5.41)(5.42)}{\leq} \|W(\omega\tau) v(\tau)\|_{H_x^s} + \varepsilon\gamma^{-1} C(s) \|B W(\omega t) v(t)\|_{H_x^{s+1}}$$

$$\overset{(5.41)}{=} \|W(\omega\tau) v(\tau)\|_{H_x^s} + \varepsilon\gamma^{-1} C(s) \|W(\omega\tau) v(\tau)\|_{H_x^{s+1}}$$

$$\overset{(5.40)}{\leq} \|v(\tau)\|_{H_x^s} + \|(W(\omega\tau) - I) v(\tau)\|_{H_x^s} + \varepsilon\gamma^{-1} C(s) \|v(\tau)\|_{H_x^{s+1}}$$

$$\overset{(5.42)}{\leq} \|v(\tau)\|_{H_x^s} + \varepsilon\gamma^{-1} C(s) \|v(\tau)\|_{H_x^{s+1}} \overset{(1.21)}{=} \|v(\tau_0)\|_{H_x^s} + \varepsilon\gamma^{-1} C(s) \|v(\tau_0)\|_{H_x^{s+1}}$$

$$\overset{(5.39)}{=} \|W^{-1}(\omega\tau_0) B^{-1} \mathcal{A}^{-1}(\omega\tau_0) h(\tau_0)\|_{H_x^s} + \varepsilon\gamma^{-1} C(s) \|W^{-1}(\omega\tau_0) B^{-1} \mathcal{A}^{-1}(\omega\tau_0) h(\tau_0)\|_{H_x^{s+1}} \,.$$

Applying the same chain of inequalities at $\tau = \tau_0$, $t = 0$, we get that the last term is

$$\leq \|h(0)\|_{H_x^s} + \varepsilon\gamma^{-1}C(s)\|h(0)\|_{H_x^{s+1}},$$

proving the second inequality in (1.23) with $\mathtt{a} := 1 - a$. The first one follows similarly.

6 Appendix A. General tame and Lipschitz estimates

In this Appendix we present standard tame and Lipschitz estimates for composition of functions and changes of variables which are used in the paper. Similar material is contained in [26], [27], [7], [2].

We first remind classical embedding, algebra, interpolation and tame estimates in the Sobolev spaces $H^s := H^s(\mathbb{T}^d, \mathbb{C})$ and $W^{s,\infty} := W^{s,\infty}(\mathbb{T}^d, \mathbb{C})$, $d \geq 1$.

Lemma 6.1. *Let $s_0 > d/2$. Then*
*(i) **Embedding.** $\|u\|_{L^\infty} \leq C(s_0)\|u\|_{s_0}$ for all $u \in H^{s_0}$.*
*(ii) **Algebra.** $\|uv\|_{s_0} \leq C(s_0)\|u\|_{s_0}\|v\|_{s_0}$ for all $u, v \in H^{s_0}$.*
*(iii) **Interpolation.** For $0 \leq s_1 \leq s \leq s_2$, $s = \lambda s_1 + (1-\lambda)s_2$,*

$$\|u\|_s \leq \|u\|_{s_1}^\lambda \|u\|_{s_2}^{1-\lambda}, \quad \forall u \in H^{s_2}. \tag{6.1}$$

Let $a_0, b_0 \geq 0$ and $p, q > 0$. For all $u \in H^{a_0+p+q}$, $v \in H^{b_0+p+q}$,

$$\|u\|_{a_0+p}\|v\|_{b_0+q} \leq \|u\|_{a_0+p+q}\|v\|_{b_0} + \|u\|_{a_0}\|v\|_{b_0+p+q}. \tag{6.2}$$

Similarly, for the $|u|_{s,\infty} := \sum_{|\beta| \leq s} |D^\beta u|_{L^\infty}$ norm,

$$|u|_{s,\infty} \leq C(s_1, s_2)|u|_{s_1,\infty}^\lambda |u|_{s_2,\infty}^{1-\lambda}, \quad \forall u \in W^{s_2,\infty}. \tag{6.3}$$

and $\forall u \in W^{a_0+p+q,\infty}$, $v \in W^{b_0+p+q,\infty}$,

$$|u|_{a_0+p,\infty}|v|_{b_0+q,\infty} \leq C(a_0, b_0, p, q)\big(|u|_{a_0+p+q,\infty}|v|_{b_0,\infty} + |u|_{a_0,\infty}|v|_{b_0+p+q,\infty}\big). \tag{6.4}$$

*(iv) **Asymmetric tame product.** For $s \geq s_0$,*

$$\|uv\|_s \leq C(s_0)\|u\|_s\|v\|_{s_0} + C(s)\|u\|_{s_0}\|v\|_s, \quad \forall u, v \in H^s. \tag{6.5}$$

*(v) **Asymmetric tame product in $W^{s,\infty}$.** For $s \geq 0$, $s \in \mathbb{N}$,*

$$|uv|_{s,\infty} \leq \tfrac{3}{2}|u|_{L^\infty}|v|_{s,\infty} + C(s)|u|_{s,\infty}|v|_{L^\infty}, \quad \forall u, v \in W^{s,\infty}. \tag{6.6}$$

*(vi) **Mixed norms asymmetric tame product.** For $s \geq 0$, $s \in \mathbb{N}$,*

$$\|uv\|_s \leq \tfrac{3}{2}|u|_{L^\infty}\|v\|_s + C(s)|u|_{s,\infty}\|v\|_0, \quad \forall u \in W^{s,\infty}, \ v \in H^s. \tag{6.7}$$

If $u := u(\lambda)$ and $v := v(\lambda)$ depend in a lipschitz way on $\lambda \in \Lambda \subset \mathbb{R}$, all the previous statements hold if we replace the norms $\|\cdot\|_s$, $|\cdot|_{s,\infty}$ with the norms $\|\cdot\|_s^{\mathrm{Lip}(\gamma)}$, $|\cdot|_{s,\infty}^{\mathrm{Lip}(\gamma)}$.

PROOF. The interpolation estimate (6.1) for the Sobolev norm (1.5) follows by Hölder inequality, see also [38], page 269. Let us prove (6.2). Let $a = a_0\lambda + a_1(1-\lambda)$, $b = b_0(1-\lambda) + b_1\lambda$, $\lambda \in [0, 1]$. Then (6.1) implies

$$\|u\|_a\|v\|_b \leq \big(\|u\|_{a_0}\|v\|_{b_1}\big)^\lambda\big(\|u\|_{a_1}\|v\|_{b_0}\big)^{1-\lambda} \leq \lambda\|u\|_{a_0}\|v\|_{b_1} + (1-\lambda)\|u\|_{a_1}\|v\|_{b_0} \tag{6.8}$$

by Young inequality. Applying (6.8) with $a = a_0 + p$, $b = b_0 + q$, $a_1 = a_0 + p + q$, $b_1 = b_0 + p + q$, then $\lambda = q/(p+q)$ and we get (6.2). Also the interpolation estimates (6.3) are classical (see e.g. [26], [10]) and (6.3) implies (6.4) as above.

(iv): see the Appendix of [10]. (v): we write, in the standard multi-index notation,

$$D^\alpha(uv) = \sum_{\beta+\gamma=\alpha} C_{\beta,\gamma}(D^\beta u)(D^\gamma v) = uD^\alpha v + \sum_{\beta+\gamma=\alpha, \beta\neq 0} C_{\beta,\gamma}(D^\beta u)(D^\gamma v). \tag{6.9}$$

Using $|(D^\beta u)(D^\gamma v)|_{L^\infty} \leq |D^\beta u|_{L^\infty}|D^\gamma v|_{L^\infty} \leq |u|_{|\beta|,\infty}|v|_{|\gamma|,\infty}$, and the interpolation inequality (6.3) for every $\beta \neq 0$ with $\lambda := |\beta|/|\alpha| \in (0,1]$ (where $|\alpha| \leq s$), we get, for any $K > 0$,

$$
\begin{aligned}
C_{\beta,\gamma}|D^\beta u|_{L^\infty}|D^\gamma v|_{L^\infty} &\leq C_{\beta,\gamma}C(s)\big(|v|_{L^\infty}|u|_{s,\infty}\big)^\lambda \big(|v|_{s,\infty}|u|_{L^\infty}\big)^{1-\lambda} \\
&= \frac{C(s)}{K}\big[(KC_{\beta,\gamma})^{\frac{1}{\lambda}}|v|_{L^\infty}|u|_{s,\infty}\big]^\lambda \big(|v|_{s,\infty}|u|_{L^\infty}\big)^{1-\lambda} \\
&\leq \frac{C(s)}{K}\big\{(KC_{\beta,\gamma})^{\frac{|\alpha|}{|\beta|}}|v|_{L^\infty}|u|_{s,\infty} + |v|_{s,\infty}|u|_{L^\infty}\big\}. \tag{6.10}
\end{aligned}
$$

Then (6.6) follows by (6.9), (6.10) taking $K := K(s)$ large enough. (vi): same proof as (v), using the elementary inequality $\|(D^\beta u)(D^\gamma v)\|_0 \leq |D^\beta u|_{L^\infty}\|D^\gamma v\|_0$. ∎

We now recall classical tame estimates for composition of functions, see [38], section 2, pages 272–275, and [42]-I, Lemma 7 in the Appendix, pages 202–203.

A function $f : \mathbb{T}^d \times B_1 \to \mathbb{C}$, where $B_1 := \{y \in \mathbb{R}^m : |y| < 1\}$, induces the composition operator

$$\tilde{f}(u)(x) := f(x, u(x), Du(x), \ldots, D^p u(x)) \tag{6.11}$$

where $D^k u(x)$ denotes the partial derivatives $\partial_x^\alpha u(x)$ of order $|\alpha| = k$ (the number m of y-variables depends on p, d).

Lemma 6.2. (Composition of functions) *Assume $f \in C^r(\mathbb{T}^d \times B_1)$. Then*
(i) For all $u \in H^{r+p}$ such that $|u|_{p,\infty} < 1$, the composition operator (6.11) is well defined and

$$\|\tilde{f}(u)\|_r \leq C\|f\|_{C^r}(\|u\|_{r+p} + 1)$$

where the constant C depends on r, d, p. If $f \in C^{r+2}$, then, for all $|u|_{p,\infty}, |h|_{p,\infty} < 1/2$,

$$\big\|\tilde{f}(u+h) - \tilde{f}(u)\big\|_r \leq C\|f\|_{C^{r+1}}(\|h\|_{r+p} + |h|_{p,\infty}\|u\|_{r+p}),$$
$$\big\|\tilde{f}(u+h) - \tilde{f}(u) - \tilde{f}'(u)[h]\big\|_r \leq C\|f\|_{C^{r+2}}|h|_{p,\infty}(\|h\|_{r+p} + |h|_{p,\infty}\|u\|_{r+p}).$$

(ii) The previous statement also holds replacing $\| \ \|_r$ with the norms $| \ |_{r,\infty}$.

Lemma 6.3. (Lipschitz estimate on parameters) *Let $d \in \mathbb{N}$, $d/2 < s_0 \leq s$, $p \geq 0$, $\gamma > 0$. Let F be a C^1-map satisfying the tame estimates: $\forall \|u\|_{s_0+p} \leq 1$, $h \in H^{s+p}$,*

$$\|F(u)\|_s \leq C(s)(1 + \|u\|_{s+p}), \tag{6.12}$$
$$\|\partial_u F(u)[h]\|_s \leq C(s)(\|h\|_{s+p} + \|u\|_{s+p}\|h\|_{s_0+p}). \tag{6.13}$$

For $\Lambda \subset \mathbb{R}$, let $u(\lambda)$ be a Lipschitz family of functions with $\|u\|_{s_0+p}^{\text{Lip}(\gamma)} \leq 1$ (see (2.2)). Then

$$\|F(u)\|_s^{\text{Lip}(\gamma)} \leq C(s)\big(1 + \|u\|_{s+p}^{\text{Lip}(\gamma)}\big).$$

The same statement also holds when all the norms $\| \ \|_s$ are replaced by $| \ |_{s,\infty}$.

PROOF. By (6.12) we get $\sup_\lambda \|F(u(\lambda))\|_s \leq C(s)(1 + \|u\|_{s+p}^{\text{Lip}(\gamma)})$. Then, denoting $u_1 := u(\lambda_1)$ and $h := u(\lambda_2) - u(\lambda_1)$, we have

$$
\begin{aligned}
\|F(u_2) - F(u_1)\|_s &\leq \int_0^1 \|\partial_u F(u_1 + t(u_2 - u_1))[h]\|_s \, dt \\
&\overset{(6.13)}{\leq}_s \|h\|_{s+p} + \|h\|_{s_0+p}\int_0^1 \big((1-t)\|u(\lambda_1)\|_{s+p} + t\|u(\lambda_2)\|_{s+p}\big) \, dt
\end{aligned}
$$

whence

$$\gamma \sup_{\substack{\lambda_1,\lambda_2\in\Lambda \\ \lambda_1\neq\lambda_2}} \frac{\|F(u(\lambda_1))-F(u(\lambda_2))\|_s}{|\lambda_1-\lambda_2|} \leq_s \|u\|_{s+p}^{\mathrm{Lip}(\gamma)} + \|u\|_{s_0+p}^{\mathrm{Lip}(\gamma)} \sup_{\lambda_1,\lambda_2} \big(\|u(\lambda_1)\|_{s+p}+\|u(\lambda_2)\|_{s+p}\big)$$

$$\leq_s \|u\|_{s+p}^{\mathrm{Lip}(\gamma)} + \|u\|_{s_0+p}^{\mathrm{Lip}(\gamma)}\|u\|_{s+p}^{\mathrm{Lip}(\gamma)} \leq C(s)\|u\|_{s+p}^{\mathrm{Lip}(\gamma)},$$

because $\|u\|_{s_0+p}^{\mathrm{Lip}(\gamma)} \leq 1$, and the lemma follows. ■

The next lemma is also classical, see for example [26], Appendix, and [27], Appendix G. The present version is proved in [2], adapting Lemma 2.3.6 on page 149 of [25], except for the part on the Lipschitz dependence on a parameter, which is proved here below.

Lemma 6.4. (Change of variable) *Let $p : \mathbb{R}^d \to \mathbb{R}^d$ be a 2π-periodic function in $W^{s,\infty}$, $s \geq 1$, with $|p|_{1,\infty} \leq 1/2$. Let $f(x) = x + p(x)$. Then:*
(i) f is invertible, its inverse is $f^{-1}(y) = g(y) = y + q(y)$ where q is 2π-periodic, $q \in W^{s,\infty}(\mathbb{T}^d,\mathbb{R}^d)$, and $|q|_{s,\infty} \leq C|p|_{s,\infty}$. More precisely,

$$|q|_{L^\infty} = |p|_{L^\infty}, \quad |Dq|_{L^\infty} \leq 2|Dp|_{L^\infty}, \quad |Dq|_{s-1,\infty} \leq C|Dp|_{s-1,\infty}. \tag{6.14}$$

where the constant C depends on d,s.
Moreover, suppose that $p = p_\lambda$ depends in a Lipschitz way by a parameter $\lambda \in \Lambda \subset \mathbb{R}$, and suppose, as above, that $|D_x p_\lambda|_{L^\infty} \leq 1/2$ for all λ. Then $q = q_\lambda$ is also Lipschitz in λ, and

$$|q|_{s,\infty}^{\mathrm{Lip}(\gamma)} \leq C\Big(|p|_{s,\infty}^{\mathrm{Lip}(\gamma)} + \big\{\sup_{\lambda\in\Lambda}|p_\lambda|_{s+1,\infty}\big\}|p|_{L^\infty}^{\mathrm{Lip}(\gamma)}\Big) \leq C|p|_{s+1,\infty}^{\mathrm{Lip}(\gamma)}, \tag{6.15}$$

The constant C depends on d,s (and is independent on γ).
(ii) If $u \in H^s(\mathbb{T}^d,\mathbb{C})$, then $u \circ f(x) = u(x+p(x))$ is also in H^s, and, with the same C as in (i),

$$\|u \circ f\|_s \leq C(\|u\|_s + |Dp|_{s-1,\infty}\|u\|_1), \tag{6.16}$$

$$\|u \circ f - u\|_s \leq C\big(|p|_{L^\infty}\|u\|_{s+1} + |p|_{s,\infty}\|u\|_2\big), \tag{6.17}$$

$$\|u \circ f\|_s^{\mathrm{Lip}(\gamma)} \leq C\big(\|u\|_{s+1}^{\mathrm{Lip}(\gamma)} + |p|_{s,\infty}^{\mathrm{Lip}(\gamma)}\|u\|_2^{\mathrm{Lip}(\gamma)}\big). \tag{6.18}$$

(6.16), (6.17) (6.18) *also hold for $u \circ g$.*
(iii) Part (ii) also holds with $\|\cdot\|_k$ replaced by $|\cdot|_{k,\infty}$, and $\|\cdot\|_s^{\mathrm{Lip}(\gamma)}$ replaced by $|\cdot|_{s,\infty}^{\mathrm{Lip}(\gamma)}$, namely

$$|u \circ f|_{s,\infty} \leq C(|u|_{s,\infty} + |Dp|_{s-1,\infty}|u|_{1,\infty}), \tag{6.19}$$

$$|u \circ f|_{s,\infty}^{\mathrm{Lip}(\gamma)} \leq C(|u|_{s+1,\infty}^{\mathrm{Lip}(\gamma)} + |Dp|_{s-1,\infty}^{\mathrm{Lip}(\gamma)}|u|_{2,\infty}^{\mathrm{Lip}(\gamma)}). \tag{6.20}$$

PROOF. The bounds (6.14), (6.16) and (6.19) are proved in [2], Appendix B. Let us prove (6.15). Denote $p_\lambda(x) := p(\lambda,x)$, and similarly for $q_\lambda, g_\lambda, f_\lambda$. Since $y = f_\lambda(x) = x + p_\lambda(x)$ if and only if $x = g_\lambda(y) = y + q_\lambda(y)$, one has

$$q_\lambda(y) + p_\lambda(g_\lambda(y)) = 0, \quad \forall \lambda \in \Lambda, \ y \in \mathbb{T}^d. \tag{6.21}$$

Let $\lambda_1,\lambda_2 \in \Lambda$, and denote, in short, $q_1 = q_{\lambda_1}$, $q_2 = q_{\lambda_2}$, and so on. By (6.21),

$$q_1 - q_2 = p_2 \circ g_2 - p_1 \circ g_1 = (p_2 \circ g_2 - p_1 \circ g_2) + (p_1 \circ g_2 - p_1 \circ g_1)$$

$$= A_2^{-1}(p_2 - p_1) + \int_0^1 A_t^{-1}(D_x p_1)\,dt\,(q_2 - q_1) \tag{6.22}$$

where $A_2^{-1}h := h \circ g_2$, $A_t^{-1}h := h \circ \big(g_1 + t[g_2 - g_1]\big)$, $t \in [0,1]$. By (6.22), the L^∞ norm of $(q_2 - q_1)$ satisfies

$$|q_2 - q_1|_{L^\infty} \leq |A_2^{-1}(p_2 - p_1)|_{L^\infty} + \int_0^1 |A_t^{-1}(D_x p_1)|_{L^\infty}\,dt\,|q_2 - q_1|_{L^\infty} \leq |p_2 - p_1|_{L^\infty} + \int_0^1 |D_x p_1|_{L^\infty}dt\,|q_2 - q_1|_{L^\infty}$$

42

whence, using the assumption $|D_x p_1|_{L^\infty} \le 1/2$,

$$|q_2 - q_1|_{L^\infty} \le 2|p_2 - p_1|_{L^\infty} .\tag{6.23}$$

By (6.22), using (6.6), the $W^{s,\infty}$ norm of $(q_2 - q_1)$, for $s \ge 0$, satisfies

$$|q_1 - q_2|_{s,\infty} \le |A_2^{-1}(p_2 - p_1)|_{s,\infty} + \frac{3}{2} \int_0^1 |A_t^{-1}(D_x p_1)|_{L^\infty}\, dt\, |q_2 - q_1|_{s,\infty} + C(s) \int_0^1 |A_t^{-1}(D_x p_1)|_{s,\infty}\, dt\, |q_2 - q_1|_{L^\infty}.$$

Since $|A_t^{-1}(D_x p_1)|_{L^\infty} = |D_x p_1|_{L^\infty} \le 1/2$,

$$\left(1 - \frac{3}{4}\right)|q_2 - q_1|_{s,\infty} \le |A_2^{-1}(p_2 - p_1)|_{s,\infty} + C(s) \int_0^1 |A_t^{-1}(D_x p_1)|_{s,\infty}\, dt\, |q_2 - q_1|_{L^\infty}.$$

Using (6.23), (6.19), (6.4) and (6.14),

$$|q_2 - q_1|_{s,\infty} \le C(s)\Big(|p_2 - p_1|_{s,\infty} + \big\{ \sup_{\lambda \in \Lambda} |p_\lambda|_{s+1,\infty} \big\}|p_2 - p_1|_{L^\infty}\Big)$$

and (6.15) follows.

Proof of (6.17). We have $u \circ f - u = \int_0^1 A_t(D_x u)\, dt\, p$ where $A_t u(x) := u(x + tp(x))$, $t \in [0,1]$. Then, by (6.7) and (6.16),

$$\begin{aligned}
\Big\| \int_0^1 A_t(D_x u)\, dt\, p \Big\|_s &\le_s \int_0^1 \|A_t(D_x u)\|_s\, dt\, |p|_{L^\infty} + \int_0^1 \|A_t(D_x u)\|_0\, dt\, |p|_{s,\infty} \\
&\le_s \|u\|_{s+1}|p|_{L^\infty} + |p|_{s,\infty}|p|_{L^\infty}\|u\|_2 + |p|_{s,\infty}\|u\|_1 ,
\end{aligned}$$

which implies (6.17).

Proof of (6.18). With the same notation as above,

$$u_2 \circ f_2 - u_1 \circ f_1 = (u_2 \circ f_2 - u_2 \circ f_1) + (u_2 \circ f_1 - u_1 \circ f_1) = \int_0^1 A_t(D_x u_2)\, dt\, (f_2 - f_1) + A_1(u_2 - u_1),$$

where $A_1 h = h \circ f_1$ and $A_t h = h \circ (f_1 + t[f_2 - f_1])$. Using (6.7) and (6.16),

$$\Big\| \int_0^1 A_t(D_x u_2)\, dt\, (f_2 - f_1) \Big\|_s \le_s \left(\|D_x u_2\|_s + \big(\sup_\lambda |D_x p_\lambda|_{s-1,\infty}\big)\|D_x u_2\|_1 \right)|p_2 - p_1|_{L^\infty} + \|D_x u_2\|_0|p_2 - p_1|_{s,\infty}$$

and $\|A_1(u_2 - u_1)\|_s \le_s \|u_2 - u_1\|_s + |D_x p_1|_{s-1,\infty}\|u_2 - u_1\|_1$. Therefore

$$\begin{aligned}
\|u_2 \circ f_2 - u_1 \circ f_1\|_s &\le_s |p_2 - p_1|_{L^\infty}\Big(\sup_\lambda \|u_\lambda\|_{s+1} + \big(\sup_\lambda |p_\lambda|_{s,\infty}\big)\big(\sup_\lambda \|u_\lambda\|_2\big)\Big) \\
&\quad + |p_2 - p_1|_{s,\infty}\big(\sup_\lambda \|u_\lambda\|_1\big) + \|u_2 - u_1\|_s + \big(\sup_\lambda |p_\lambda|_{s,\infty}\big) \|u_2 - u_1\|_1
\end{aligned}$$

whence (6.18) follows. The proof of (6.20) is the same as for (6.18), replacing all norms $\|\cdot\|_s$ with $|\cdot|_{s,\infty}$. $\blacksquare$

Lemma 6.5. (Composition) *Suppose that for all* $\|u\|_{s_0+\mu_i} \le 1$ *the operator* $\mathcal{Q}_i(u)$ *satisfies*

$$\|\mathcal{Q}_i h\|_s \le C(s)\big(\|h\|_{s+\tau_i} + \|u\|_{s+\mu_i}\|h\|_{s_0+\tau_i}\big), \quad i = 1, 2.\tag{6.24}$$

Let $\tau := \max\{\tau_1, \tau_2\}$, $\mu := \max\{\mu_1, \mu_2\}$. *Then, for all*

$$\|u\|_{s_0+\tau+\mu} \le 1 ,\tag{6.25}$$

the composition operator $\mathcal{Q} := \mathcal{Q}_1 \circ \mathcal{Q}_2$ *satisfies the tame estimate*

$$\|\mathcal{Q} h\|_s \le C(s)\big(\|h\|_{s+\tau_1+\tau_2} + \|u\|_{s+\tau+\mu}\|h\|_{s_0+\tau_1+\tau_2}\big).\tag{6.26}$$

Moreover, if $\mathcal{Q}_1$, $\mathcal{Q}_2$, u *and* h *depend in a lipschitz way on a parameter* λ, *then* (6.26) *also holds with* $\|\cdot\|_s$ *replaced by* $\|\cdot\|_s^{\mathrm{Lip}(\gamma)}$.

PROOF. Apply the estimates for (6.24) to Φ_1 first, then to Φ_2, using condition (6.25). $\blacksquare$

7 Appendix B: proof of Lemmata 3.2 and 3.3

The proof is elementary. It is based on a repeated use of the tame estimates of the Lemmata of the Appendix A. For convenience, we split it into many points. We remind that $\mathfrak{s}_0 := (\nu + 2)/2$ is fixed (it plays the role of the constant s_0 in Lemma 6.1).

Estimates in Step 1.

1. — We prove that $b_3 = b$ defined in (3.17) satisfies the tame estimates

$$\|b_3 - 1\|_s \leq \varepsilon \, C(s)\big(1 + \|u\|_{s+3}\big), \tag{7.1}$$

$$\|\partial_u b_3(u)[h]\|_s \leq \varepsilon \, C(s)\big(\|h\|_{s+3} + \|u\|_{s+3}\|h\|_{\mathfrak{s}_0+3}\big), \tag{7.2}$$

$$\|b_3 - 1\|_s^{\mathrm{Lip}(\gamma)} \leq \varepsilon \, C(s)\big(1 + \|u\|_{s+3}^{\mathrm{Lip}(\gamma)}\big). \tag{7.3}$$

Proof of (7.1). Write $b_3 = b$ (see (3.17)) as

$$b_3 - 1 = \psi\big(M[g(a_3) - g(0)]\big) - \psi(0), \quad \psi(t) := (1+t)^{-3}, \quad Mh := \frac{1}{2\pi} \int_{\mathbb{T}} h \, dx, \quad g(t) := (1+t)^{-\frac{1}{3}}. \tag{7.4}$$

Thus, for ε small,

$$\|b_3 - 1\|_s \leq C(s)\|M[g(a_3) - g(0)]\|_s \leq C(s)\|g(a_3) - g(0)\|_s \leq C(s)\|a_3\|_s.$$

In the first inequality we have applied Lemma 6.2(i) to the function ψ, with $u = 0$, $p = 0$, $h = M[g(a_3) - g(0)]$. In the second inequality we have used the trivial fact that $\|Mh\|_s \leq \|h\|_s$ for all h. In the third inequality we have applied again Lemma 6.2(i) to the function g, with $u = 0$, $p = 0$, $h = a_3$. Finally we estimate a_3 by (3.4) with $s_0 = \mathfrak{s}_0$, which holds for $s + 2 \leq q$.

Proof of (7.2). Using (7.4), the derivative of b_3 with respect to u in the direction h is

$$\partial_u b_3(u)[h] = \psi'\big(M[g(a_3) - g(0)]\big) M\big(g'(a_3)\partial_u a_3[h]\big).$$

Then use (6.5), Lemma 6.2(i) applied to the functions ψ' and g', and (3.5).

Proof of (7.3). It follows from (7.1), (7.2) and Lemma 6.3.

2. — Using the definition (3.16) of ρ_0, estimates (7.1), (7.2), (7.3) for b_3 and estimates (3.4), (3.5), (3.6) for a_3, one proves that ρ_0 also satisfies the same estimates (7.1), (7.2), (7.3) as $(b_3 - 1)$. Since $\beta = \partial_x^{-1}\rho_0$ (see (3.18)), by Lemma 6.1(i) we get

$$|\beta|_{s,\infty} \leq C(s)\|\beta\|_{s+\mathfrak{s}_0} \leq C(s)\|\rho_0\|_{s+\mathfrak{s}_0} \leq \varepsilon \, C(s)\big(1 + \|u\|_{s+\mathfrak{s}_0+3}\big), \tag{7.5}$$

and, with the same chain of inequalities,

$$|\partial_u\beta(u)[h]|_{s,\infty} \leq \varepsilon \, C(s)\big(\|h\|_{s+\mathfrak{s}_0+3} + \|u\|_{s+\mathfrak{s}_0+3}\|h\|_{\mathfrak{s}_0+3}\big). \tag{7.6}$$

Then Lemma 6.3 implies

$$|\beta|_{s,\infty}^{\mathrm{Lip}(\gamma)} \leq \varepsilon \, C(s)\big(1 + \|u\|_{s+\mathfrak{s}_0+3}^{\mathrm{Lip}(\gamma)}\big), \tag{7.7}$$

for all $s + \mathfrak{s}_0 + 3 \leq q$. Note that $x \mapsto x + \beta(\varphi, x)$ is a well-defined diffeomorphism if $|\beta|_{1,\infty} \leq 1/2$, and, by (7.5), this condition is satisfied provided $\varepsilon \, C\big(1 + \|u\|_{\mathfrak{s}_0+4}\big) \leq 1/2$.

Let $(\varphi, y) \mapsto (\varphi, y + \tilde{\beta}(\varphi, y))$ be the inverse diffeomorphism of $(\varphi, x) \mapsto (\varphi, x + \beta(\varphi, x))$. By Lemma 6.4$(i)$ on the torus $\mathbb{T}^{\nu+1}$, $\tilde{\beta}$ satisfies

$$|\tilde{\beta}|_{s,\infty} \leq C|\beta|_{s,\infty} \overset{(7.5)}{\leq} \varepsilon \, C(s)\big(1 + \|u\|_{s+3+\mathfrak{s}_0}\big). \tag{7.8}$$

Writing explicitly the dependence on u, we have $\tilde{\beta}(\varphi, y; u) + \beta\big(\varphi, y + \tilde{\beta}(\varphi, y; u); u\big) = 0$. Differentiating the last equality with respect to u in the direction h gives

$$(\partial_u\tilde{\beta})[h] = -{}^{-1}\Big(\frac{\partial_u\beta[h]}{1 + \beta_x}\Big),$$

therefore, applying Lemma 6.4(iii) to deal with $\mathcal{A}^{-1}$, (6.6) for the product $(\partial_u \beta[h])(1 + \beta_x)^{-1}$, the estimates (7.5), (7.6), (7.7) for β, and (6.2) (with $a_0 = \mathfrak{s}_0 + 3$, $b_0 = \mathfrak{s}_0 + 4$, $p = 1$, $q = s - 1$), we obtain (for $s + \mathfrak{s}_0 + 4 \le q$)

$$|\partial_u \tilde{\beta}(u)[h]|_{s,\infty} \le \varepsilon\, C(s)\big(\|h\|_{s+3+\mathfrak{s}_0} + \|u\|_{s+4+\mathfrak{s}_0}\|h\|_{3+\mathfrak{s}_0}\big). \tag{7.9}$$

Then, using Lemma 6.3 with $p = 4 + \mathfrak{s}_0$, the bounds (7.8), (7.9) imply

$$|\tilde{\beta}|_{s,\infty}^{\mathrm{Lip}(\gamma)} \le \varepsilon\, C(s)\big(1 + \|u\|_{s+4+\mathfrak{s}_0}^{\mathrm{Lip}(\gamma)}\big). \tag{7.10}$$

3. — ESTIMATES OF $\mathcal{A}(u)$ AND $\mathcal{A}(u)^{-1}$. By (6.16), (7.5) and (7.8),

$$\|\mathcal{A}(u)h\|_s + \|\mathcal{A}(u)^{-1}h\|_s \le C(s)\big(\|h\|_s + \|u\|_{s+\mathfrak{s}_0+3}\|h\|_1\big). \tag{7.11}$$

Moreover, by (6.18), (7.7) and (7.10),

$$\|\mathcal{A}(u)h\|_s^{\mathrm{Lip}(\gamma)} + \|\mathcal{A}(u)^{-1}h\|_s^{\mathrm{Lip}(\gamma)} \le C(s)\big(\|h\|_{s+1}^{\mathrm{Lip}(\gamma)} + \|u\|_{s+\mathfrak{s}_0+4}^{\mathrm{Lip}(\gamma)}\|h\|_2^{\mathrm{Lip}(\gamma)}\big). \tag{7.12}$$

Since $\mathcal{A}(u)g(\varphi, x) = g(\varphi, x + \beta(\varphi, x; u))$, the derivative of $\mathcal{A}(u)g$ with respect to u in the direction h is the product $\partial_u\big(\mathcal{A}(u)g\big)[h] = (\mathcal{A}(u)g_x)\, \partial_u \beta(u)[h]$. Then, by (6.7), (7.6) and (7.11),

$$\|\partial_u(\mathcal{A}(u)g)[h]\|_s \le \varepsilon C(s)\Big(\|g\|_{s+1}\|h\|_{\mathfrak{s}_0+3} + \|g\|_2\|h\|_{s+\mathfrak{s}_0+3} + \|u\|_{s+\mathfrak{s}_0+3}\|g\|_2\|h\|_{\mathfrak{s}_0+3}\Big). \tag{7.13}$$

Similarly $\partial_u(\mathcal{A}(u)^{-1}g)[h] = (\mathcal{A}(u)^{-1}g_x)\, \partial_u \tilde{\beta}(u)[h]$, therefore (6.7), (7.9), (7.11) imply that

$$\|\partial_u(\mathcal{A}^{-1}(u)g)[h]\|_s \le \varepsilon C(s)\Big(\|g\|_{s+1}\|h\|_{\mathfrak{s}_0+3} + \|g\|_2\|h\|_{s+\mathfrak{s}_0+3} + \|u\|_{s+\mathfrak{s}_0+4}\|g\|_2\|h\|_{\mathfrak{s}_0+3}\Big). \tag{7.14}$$

4. — The coefficients b_0, b_1, b_2 are given in (3.12), (3.13). By (6.7), (7.11), (3.61), (7.5) and (3.4),

$$\|b_i\|_s \le \varepsilon C(s)(1 + \|u\|_{s+\mathfrak{s}_0+6}), \quad i = 0, 1, 2. \tag{7.15}$$

Moreover, in analogous way, by (6.7), (7.12), (3.61), (7.7) and (3.6),

$$\|b_i\|_s^{\mathrm{Lip}(\gamma)} \le \varepsilon C(s)(1 + \|u\|_{s+\mathfrak{s}_0+7}^{\mathrm{Lip}(\gamma)}), \quad i = 0, 1, 2. \tag{7.16}$$

Now we estimate the derivative with respect to u of b_1. The estimates for b_0 and b_2 are analogous. By (3.12) we write $b_1(u) = \mathcal{A}(u)^{-1}b_1^*(u)$ where $b_1^* := \omega \cdot \partial_\varphi \beta + (1 + a_3)\beta_{xxx} + a_2\beta_{xx} + a_1(1 + \beta_x)$. The bounds (3.5), (7.6), (7.5), (3.61), and (6.7) imply that

$$\|\partial_u b_1^*(u)[h]\|_s \le \varepsilon C(s)\big(\|h\|_{s+\mathfrak{s}_0+6} + \|u\|_{s+\mathfrak{s}_0+6}\|h\|_{\mathfrak{s}_0+6}\big). \tag{7.17}$$

Now,

$$\partial_u b_1(u)[h] = \partial_u\big(\mathcal{A}(u)^{-1}b_1^*(u)\big)[h] = (\partial_u \mathcal{A}(u)^{-1})(b_1^*(u))[h] + \mathcal{A}(u)^{-1}(\partial_u b_1^*(u)[h]). \tag{7.18}$$

Then (6.5), (7.18), (7.11), (7.14), (6.2) (with $a_0 = \mathfrak{s}_0 + 4$, $\beta_0 = \mathfrak{s}_0 + 6$, $p = s - 1$, $q = 1$) (7.17) imply

$$\|\partial_u \mathcal{A}(u)^{-1}(b_1^*(u))[h]\|_s \le \varepsilon C(s)\big(\|h\|_{s+\mathfrak{s}_0+3} + \|u\|_{s+\mathfrak{s}_0+7}\|h\|_{\mathfrak{s}_0+3}\big) \tag{7.19}$$

$$\|\mathcal{A}(u)^{-1}\partial_u b_1^*(u)[h]\|_s \le \varepsilon C(s)\big(\|h\|_{s+\mathfrak{s}_0+6} + \|u\|_{s+\mathfrak{s}_0+6}\|h\|_{\mathfrak{s}_0+6}\big). \tag{7.20}$$

Finally (7.18), (7.19) and (7.20) imply

$$\|\partial_u b_1(u)[h]\|_s \le \varepsilon C(s)\big(\|h\|_{s+\mathfrak{s}_0+6} + \|u\|_{s+\mathfrak{s}_0+7}\|h\|_{\mathfrak{s}_0+6}\big), \tag{7.21}$$

which holds for all $s + \mathfrak{s}_0 + 7 \le q$.

Estimates in Step 2.

5. — We prove that the coefficient m_3, defined in (3.30), satisfies the following estimates:

$$|m_3 - 1|, |m_3 - 1|^{\text{Lip}(\gamma)} \leq \varepsilon C \tag{7.22}$$

$$|\partial_u m_3(u)[h]| \leq \varepsilon C \|h\|_{\mathfrak{s}_0+3}. \tag{7.23}$$

Using (3.30) (7.1), (3.61)

$$|m_3 - 1| \leq \frac{1}{(2\pi)^\nu} \int_{\mathbb{T}^\nu} |b_3 - 1| \, d\varphi \leq C \|b_3 - 1\|_{\mathfrak{s}_0} \leq \varepsilon C.$$

Similarly we get the Lipschitz part of (7.22). The estimate (7.23) follows by (7.2), since

$$|\partial_u m_3(u)[h]| \leq \frac{1}{(2\pi)^\nu} \int_{\mathbb{T}^\nu} |\partial_u b_3(u)[h]| \, d\varphi \leq C \|\partial_u b_3(u)[h]\|_{\mathfrak{s}_0} \leq \varepsilon C \|h\|_{\mathfrak{s}_0+3}.$$

6. — ESTIMATES OF α. The function $\alpha(\varphi)$, defined in (3.31), satisfies

$$|\alpha|_{s,\infty} \leq \varepsilon \gamma_0^{-1} C(s) \big(1 + \|u\|_{s+\tau_0+\mathfrak{s}_0+3}\big) \tag{7.24}$$

$$|\alpha|_{s,\infty}^{\text{Lip}(\gamma)} \leq \varepsilon \gamma_0^{-1} C(s) \big(1 + \|u\|_{s+\tau_0+\mathfrak{s}_0+3}^{\text{Lip}(\gamma)}\big) \tag{7.25}$$

$$|\partial_u \alpha(u)[h]|_{s,\infty} \leq \varepsilon \gamma_0^{-1} C(s) \big(\|h\|_{s+\tau_0+\mathfrak{s}_0+3} + \|u\|_{s+\tau_0+\mathfrak{s}_0+3} \|h\|_{\mathfrak{s}_0+3}\big). \tag{7.26}$$

Remember that $\omega = \lambda \bar{\omega}$, and $|\bar{\omega} \cdot l| \geq 3\gamma_0 |l|^{-\tau_0}$, $\forall l \neq 0$, see (1.2). By (7.1) and (7.22),

$$|\alpha|_{s,\infty} \leq \|\alpha\|_{s+\mathfrak{s}_0} \leq C\gamma_0^{-1} \|b_3 - m_3\|_{s+\mathfrak{s}_0+\tau_0} \leq C(s)\gamma_0^{-1}\varepsilon(1 + \|u\|_{s+\tau_0+\mathfrak{s}_0+3})$$

proving (7.24). Then (7.25) holds similarly using (7.3) and $(\omega \cdot \partial_\varphi)^{-1} = \lambda^{-1} (\bar{\omega} \cdot \partial_\varphi)^{-1}$. Differentiating formula (3.31) with respect to u in the direction h gives

$$\partial_u \alpha(u)[h] = (\lambda \bar{\omega} \cdot \partial_\varphi)^{-1} \Big(\frac{\partial_u b_3(u)[h] m_3 - b_3 \partial_u m_3(u)[h]}{m_3^2} \Big)$$

then, the standard Sobolev embedding, (7.1), (7.2), (7.22), (7.23) imply (7.26). Estimates (7.25) and (7.26) hold for $s + \tau_0 + \mathfrak{s}_0 + 3 \leq q$. Note that (3.23) is a well-defined diffeomorphism if $|\alpha|_{1,\infty} \leq 1/2$, and, by (7.25), this holds by (3.59).

7. — ESTIMATES OF $\tilde{\alpha}$. Let $\vartheta \to \vartheta + \omega \tilde{\alpha}(\vartheta)$ be the inverse change of variable of (3.23). The following estimates hold:

$$|\tilde{\alpha}|_{s,\infty} \leq \varepsilon \gamma_0^{-1} C(s) \big(1 + \|u\|_{s+\tau_0+\mathfrak{s}_0+3}\big) \tag{7.27}$$

$$|\tilde{\alpha}|_{s,\infty}^{\text{Lip}(\gamma)} \leq \varepsilon \gamma_0^{-1} C(s) \big(1 + \|u\|_{s+\tau_0+\mathfrak{s}_0+4}^{\text{Lip}(\gamma)}\big) \tag{7.28}$$

$$|\partial_u \tilde{\alpha}(u)[h]|_{s,\infty} \leq \varepsilon \gamma_0^{-1} C(s) \big(\|h\|_{s+\tau_0+\mathfrak{s}_0+3} + \|u\|_{s+\tau_0+\mathfrak{s}_0+4} \|h\|_{\tau_0+\mathfrak{s}_0+3}\big). \tag{7.29}$$

The bounds (7.27), (7.28) follow by (6.14), (7.24), and (6.15), (7.25), respectively. To estimate the partial derivative of $\tilde{\alpha}$ with respect to u we differentiate the identity $\tilde{\alpha}(\vartheta; u) + \alpha(\vartheta + \omega\tilde{\alpha}(\vartheta; u); u) = 0$, which gives

$$\partial_u \tilde{\alpha}(u)[h] = -B^{-1} \Big(\frac{\partial_u \alpha[h]}{1 + \omega \cdot \partial_\varphi \alpha} \Big).$$

Then applying Lemma 6.4(iii) to deal with B^{-1}, (6.6) for the product $\partial_u \alpha[h] (1 + \omega \cdot \partial_\varphi \alpha)^{-1}$, and estimates (7.25), (7.26), (6.2), we obtain (7.29).

8. — The transformations $B(u)$ and $B(u)^{-1}$, defined in (3.24) resp. (3.25), satisfy the following estimates:

$$\|B(u)h\|_s + \|B(u)^{-1}h\|_s \leq C(s)\big(\|h\|_s + \|u\|_{s+\tau_0+\mathfrak{s}_0+3}\|h\|_1\big) \tag{7.30}$$

$$\|B(u)h\|_s^{\text{Lip}(\gamma)} + \|B(u)^{-1}h\|_s^{\text{Lip}(\gamma)} \leq C(s)\big(\|h\|_{s+1}^{\text{Lip}(\gamma)} + \|u\|_{s+\tau_0+\mathfrak{s}_0+4}^{\text{Lip}(\gamma)}\|h\|_2^{\text{Lip}(\gamma)}\big) \tag{7.31}$$

$$\|\partial_u(B(u)g)[h]\|_s \leq C(s)\big(\|g\|_{s+1}\|h\|_{\sigma_0} + \|g\|_1\|h\|_{s+\sigma_0} + \|u\|_{s+\sigma_0}\|g\|_2\|h\|_{\sigma_0}\big) \tag{7.32}$$

$$\|\partial_u(B(u)^{-1}g)[h]\|_s \leq C(s)\big(\|g\|_{s+1}\|h\|_{\sigma_0} + \|g\|_1\|h\|_{s+\sigma_0} + \|u\|_{s+\sigma_0+1}\|g\|_2\|h\|_{\sigma_0}\big) \tag{7.33}$$

where $\sigma_0 := \tau_0 + \mathfrak{s}_0 + 3$. Estimates (7.30) and (7.31) follow by Lemma 6.4(ii) and (7.24), (7.27), (7.25), (7.28). The derivative of $B(u)g$ with respect to u in the direction h is the product fz where $f := B(u)(\omega \cdot \partial_\varphi g)$ and $z := \partial_u \alpha(u)[h]$. By (6.7), $\|fz\|_s \leq C(s)(\|f\|_s |z|_{L^\infty} + \|f\|_0 |z|_{s,\infty})$. Then (7.26), (7.30) imply (7.32). In analogous way, (7.29) and (7.30) give (7.33).

9. — ESTIMATES OF ρ. The function ρ defined in (3.26), namely $\rho = 1 + B^{-1}(\omega \cdot \partial_\varphi \alpha)$, satisfies

$$|\rho - 1|_{s,\infty} \leq \varepsilon \gamma_0^{-1} C(s)(1 + \|u\|_{s+\tau_0+\mathfrak{s}_0+4}) \tag{7.34}$$

$$|\rho - 1|_{s,\infty}^{\mathrm{Lip}(\gamma)} \leq \varepsilon \gamma_0^{-1} C(s)(1 + \|u\|_{s+\tau_0+\mathfrak{s}_0+5}^{\mathrm{Lip}(\gamma)}) \tag{7.35}$$

$$\|\partial_u \rho(u)[h]\|_s \leq \varepsilon \gamma_0^{-1} C(s)\big(\|h\|_{s+\tau_0+\mathfrak{s}_0+4} + \|u\|_{s+\tau_0+\mathfrak{s}_0+5}\|h\|_{\tau_0+\mathfrak{s}_0+4}\big). \tag{7.36}$$

The bound (7.34) follows by (3.26), (6.19), (7.24), (3.59). Similarly (7.35) follows by (6.20), (7.25) and (3.61). Differentiating (3.26) with respect to u in the direction h we obtain

$$\partial_u \rho(u)[h] = \partial_u B(u)^{-1}(\omega \cdot \partial_\varphi \alpha)[h] + B(u)^{-1}\big(\omega \cdot \partial_\varphi (\partial_u \alpha(u)[h])\big).$$

By (7.33), (7.24), and (3.59), we get

$$\|\partial_u B(u)^{-1}(\omega \cdot \partial_\varphi \alpha)[h]\|_s \leq \varepsilon \gamma_0^{-1} C(s)\big(\|h\|_{s+\tau_0+\mathfrak{s}_0+3} + \|u\|_{s+\tau_0+\mathfrak{s}_0+5}\|h\|_{\tau_0+\mathfrak{s}_0+3}\big). \tag{7.37}$$

Using (7.30), (7.26), (3.59), and applying (6.2), one has

$$\|B(u)^{-1}\big(\omega \cdot \partial_\varphi(\partial_u \alpha(u)[h])\big)\|_s \leq \varepsilon \gamma_0^{-1} C(s)\big(\|h\|_{s+\tau_0+\mathfrak{s}_0+4} + \|u\|_{s+\tau_0+\mathfrak{s}_0+4}\|h\|_{\tau_0+\mathfrak{s}_0+4}\big). \tag{7.38}$$

Then (7.37) and (7.38) imply (7.36), for all $s + \tau_0 + \mathfrak{s}_0 + 5 \leq q$.

10. — The coefficients c_0, c_1, c_2 defined in (3.33) satisfy the following estimates: for $i = 0, 1, 2$, $s \geq \mathfrak{s}_0$,

$$\|c_i\|_s \leq \varepsilon C(s)\big(1 + \|u\|_{s+\tau_0+\mathfrak{s}_0+6}\big), \tag{7.39}$$

$$\|c_i\|_s^{\mathrm{Lip}(\gamma)} \leq \varepsilon C(s)\big(1 + \|u\|_{s+\tau_0+\mathfrak{s}_0+7}^{\mathrm{Lip}(\gamma)}\big), \tag{7.40}$$

$$\|\partial_u c_i[h]\|_s \leq \varepsilon C(s)\big(\|h\|_{s+\tau_0+\mathfrak{s}_0+6} + \|u\|_{s+\tau_0+\mathfrak{s}_0+7}\|h\|_{\tau_0+2\mathfrak{s}_0+6}\big). \tag{7.41}$$

The definition of c_i in (3.33), (6.7), (3.59), (7.30), (7.34), (7.15) and $\varepsilon \gamma_0^{-1} < 1$, imply (7.39). Similarly (3.61), (7.31), (7.35) and (7.16) imply (7.40). Finally (7.41) follows from differentiating the formula of $c_i(u)$ and using (3.59), (7.15), (7.33), (7.30), (6.5)-(6.7), (7.34), (7.36).

Estimates in the step 3.

11. — The function v defined in (3.38) satisfies the following estimates:

$$\|v - 1\|_s \leq \varepsilon C(s)\big(1 + \|u\|_{s+\tau_0+\mathfrak{s}_0+6}\big) \tag{7.42}$$

$$\|v - 1\|_s^{\mathrm{Lip}(\gamma)} \leq \varepsilon C(s)\big(1 + \|u\|_{s+\tau_0+\mathfrak{s}_0+7}^{\mathrm{Lip}(\gamma)}\big) \tag{7.43}$$

$$\|\partial_u v[h]\|_s \leq \varepsilon C(s)\big(\|h\|_{s+\tau_0+\mathfrak{s}_0+6} + \|u\|_{s+\tau_0+\mathfrak{s}_0+7}\|h\|_{\tau_0+2\mathfrak{s}_0+6}\big) \tag{7.44}$$

In order to prove (7.42) we apply the Lemma 6.2(i) with $f(t) := \exp(t)$ (and $u = 0$, $p = 0$):

$$\|v - 1\|_s = \left\| f\Big(-\frac{\partial_y^{-1} c_2}{3m_3}\Big) - f(0) \right\|_s \overset{(7.22)}{\leq} C\|c_2\|_s \overset{(7.39)}{\leq} \varepsilon C(s)\big(1 + \|u\|_{s+\tau_0+\mathfrak{s}_0+6}\big).$$

Similarly (7.43) follows. Differentiating formula (3.38) we get

$$\partial_u v[h] = -f'\Big(-\frac{\partial_y^{-1} c_2}{3m_3}\Big)\left\{ \frac{1}{3m_3}\partial_u\Big(\partial_y^{-1} c_2\Big)[h] - \frac{\partial_y^{-1} c_2 \partial_u m_3[h]}{3m_3^2} \right\}.$$

Then using (3.59), (6.5), Lemma 6.2(i) applied to $f' = f$, and the estimates (7.39), (7.41), (7.22) and (7.23) we get (7.44).

12. — The multiplication operator $\mathcal{M}$ defined in (3.34) and its inverse $\mathcal{M}^{-1}$ (which is the multiplication operator by v^{-1}) both satisfy

$$\|\mathcal{M}^{\pm 1} h\|_s \leq C(s)\big(\|h\|_s + \|u\|_{s+\tilde{\sigma}}\|h\|_{\mathfrak{s}_0}\big), \tag{7.45}$$

$$\|\mathcal{M}^{\pm 1} h\|_s^{\mathrm{Lip}(\gamma)} \leq C(s)\big(\|h\|_s^{\mathrm{Lip}(\gamma)} + \|u\|_{s+\tilde{\sigma}+1}^{\mathrm{Lip}(\gamma)}\|h\|_{\mathfrak{s}_0}^{\mathrm{Lip}(\gamma)}\big), \tag{7.46}$$

$$\|\partial_u \mathcal{M}^{\pm 1}(u)g[h]\|_s \leq \varepsilon C(s)\big(\|g\|_s\|h\|_{\mathfrak{s}_0+\tilde{\sigma}} + \|g\|_{\mathfrak{s}_0}\|h\|_{s+\tilde{\sigma}} + \|u\|_{s+\tilde{\sigma}+1}\|g\|_{\mathfrak{s}_0}\|h\|_{\mathfrak{s}_0+\tilde{\sigma}}\big), \tag{7.47}$$

with $\tilde{\sigma} := \tau_0 + \mathfrak{s}_0 + 6$.

The inequalities (7.45)-(7.47) follow by (3.59), (3.61), (6.5), (7.42)-(7.44).

13. — The coefficients d_1, d_0, defined in (3.41), satisfy, for $i = 0, 1$

$$\|d_i\|_s \leq \varepsilon C(s)(1 + \|u\|_{s+\tau_0+\mathfrak{s}_0+9}), \tag{7.48}$$

$$\|d_i\|_s^{\mathrm{Lip}(\gamma)} \leq \varepsilon C(s)(1 + \|u\|_{s+\tau_0+\mathfrak{s}_0+10}^{\mathrm{Lip}(\gamma)}), \tag{7.49}$$

$$\|\partial_u d_i(u)[h]\|_s \leq \varepsilon C(s)\big(\|h\|_{s+\tau_0+\mathfrak{s}_0+9} + \|u\|_{s+\tau_0+\mathfrak{s}_0+10}\|h\|_{\tau_0+2\mathfrak{s}_0+9}\big), \tag{7.50}$$

by (6.5), (3.59), (3.61), (7.39)-(7.41) and (7.42)-(7.44).

Estimates in the Step 4.

14. — The constant m_1 defined in (3.46) satisfies

$$|m_1| + |m_1|^{\mathrm{Lip}(\gamma)} \leq \varepsilon C, \quad |\partial_u m_1(u)[h]| \leq \varepsilon C\|h\|_{\tau_0+2\mathfrak{s}_0+9}, \tag{7.51}$$

by (3.61), (7.48)-(7.50).

15. — The function $p(\vartheta)$ defined in (3.47) satisfies the following estimates:

$$|p|_{s,\infty} \leq \varepsilon\gamma_0^{-1}C(s)(1 + \|u\|_{s+2\tau_0+2\mathfrak{s}_0+9}) \tag{7.52}$$

$$|p|_{s,\infty}^{\mathrm{Lip}(\gamma)} \leq \varepsilon\gamma_0^{-1}C(s)(1 + \|u\|_{s+2\tau_0+2\mathfrak{s}_0+10}^{\mathrm{Lip}(\gamma)}) \tag{7.53}$$

$$|\partial_u p(u)[h]|_{s,\infty} \leq \varepsilon\gamma_0^{-1}C(s)\big(\|h\|_{s+2\tau_0+2\mathfrak{s}_0+9} + \|u\|_{s+2\tau_0+2\mathfrak{s}_0+10}\|h\|_{\tau_0+2\mathfrak{s}_0+9}\big). \tag{7.54}$$

which follow by (7.48)-(7.50) and (7.51) applying the same argument used in the proof of (7.25).

16. — The operators $\mathcal{T}$, $\mathcal{T}^{-1}$ defined in (3.42) satisfy

$$\|\mathcal{T}^{\pm 1} h\|_s \leq C(s)\big(\|h\|_s + \|u\|_{s+\bar{\sigma}}\|h\|_1\big) \tag{7.55}$$

$$\|\mathcal{T}^{\pm 1} h\|_s^{\mathrm{Lip}(\gamma)} \leq C(s)\big(\|h\|_{s+1}^{\mathrm{Lip}(\gamma)} + \|u\|_{s+\bar{\sigma}+1}^{\mathrm{Lip}(\gamma)}\|h\|_2^{\mathrm{Lip}(\gamma)}\big) \tag{7.56}$$

$$\|\partial_u(\mathcal{T}^{\pm 1}(u)g)[h]\|_s \leq \varepsilon\gamma_0^{-1}C(s)\big(\|g\|_{s+1}\|h\|_{\bar{\sigma}} + \|g\|_1\|h\|_{s+\bar{\sigma}} + \|u\|_{s+\bar{\sigma}+1}\|g\|_2\|h\|_{\bar{\sigma}}\big), \tag{7.57}$$

with $\bar{\sigma} := 2\tau_0 + 2\mathfrak{s}_0 + 9$. The estimates (7.55) and (7.56) follow by (6.16), (6.18) and using (7.52) and (7.53). The derivative $\partial_u(\mathcal{T}(u)g)[h]$ is the product $(\mathcal{T}(u)g_y)\,\partial_u p(u)[h]$. Hence (6.7), (7.55) and (7.54) imply (7.57).

17. — The coefficients e_0, e_1, defined in (3.43), satisfy the following estimates: for $i = 0, 1$

$$\|e_i\|_s \leq \varepsilon C(s)(1 + \|u\|_{s+2\tau_0+2\mathfrak{s}_0+9}), \tag{7.58}$$

$$\|e_i\|_s^{\mathrm{Lip}(\gamma)} \leq \varepsilon C(s)(1 + \|u\|_{s+2\tau_0+2\mathfrak{s}_0+10}^{\mathrm{Lip}(\gamma)}), \tag{7.59}$$

$$\|\partial_u e_i(u)[h]\|_s \leq \varepsilon C(s)\big(\|h\|_{s+2\tau_0+2\mathfrak{s}_0+9} + \|u\|_{s+2\tau_0+2\mathfrak{s}_0+10}\|h\|_{2\tau_0+2\mathfrak{s}_0+9}\big). \tag{7.60}$$

The estimates (7.58), (7.59) follow by (3.59), (3.61), (3.45), (7.48), (7.49), (7.55) and (7.56). The estimate (7.60) follows differentiating the formulae of e_0 and e_1 in (3.43), and applying (7.48), (7.50), (7.55) and (7.57).

Estimates in the Step 5.

18. — The function w defined in (3.54) satisfies the following estimates:

$$\|w\|_s \leq \varepsilon C(s)(1 + \|u\|_{s+2\tau_0+2\mathfrak{s}_0+9}) \tag{7.61}$$

$$\|w\|_s^{\mathrm{Lip}(\gamma)} \leq \varepsilon C(s)(1 + \|u\|_{s+2\tau_0+2\mathfrak{s}_0+10}^{\mathrm{Lip}(\gamma)}) \tag{7.62}$$

$$\|\partial_u w(u)[h]\|_s \leq \varepsilon C(s)\big(\|h\|_{s+2\tau_0+2\mathfrak{s}_0+9} + \|u\|_{s+2\tau_0+2\mathfrak{s}_0+10}\|h\|_{2\tau_0+2\mathfrak{s}_0+9}\big) \tag{7.63}$$

which follow by (7.22), (7.23), (7.51), (7.58)-(7.60), (3.59), (3.61).

19. — The operator $\mathcal{S} = I + w\partial_x^{-1}$, defined in (3.49), and its inverse $\mathcal{S}^{-1}$ both satisfy the following estimates (where the s-decay norm $|\cdot|_s$ is defined in (2.3)):

$$
\begin{align}
|\mathcal{S}^{\pm 1} - I|_s &\leq \varepsilon C(s)(1 + \|u\|_{s+2\tau_0+2\mathfrak{s}_0+9}), \tag{7.64} \\
|\mathcal{S}^{\pm 1} - I|_s^{\mathrm{Lip}(\gamma)} &\leq \varepsilon C(s)(1 + \|u\|_{s+2\tau_0+2\mathfrak{s}_0+10}^{\mathrm{Lip}(\gamma)}), \tag{7.65} \\
\left|\partial_u \mathcal{S}^{\pm 1}(u)[h]\right|_s &\leq \varepsilon C(s)\big(\|h\|_{s+2\tau_0+2\mathfrak{s}_0+9} + \|u\|_{s+2\tau_0+2\mathfrak{s}_0+10}\|h\|_{2\tau_0+3\mathfrak{s}_0+9}\big). \tag{7.66}
\end{align}
$$

Thus (7.64)-(7.66) for $\mathcal{S}$ follow by (7.61)-(7.63) and the fact that the matrix decay norm $|\partial_x^{-1}|_s \leq 1$, $s \geq 0$, using (2.5), (2.6), (2.8), (2.9). The operator $\mathcal{S}^{-1}$ satisfies the same bounds (7.64)-(7.65) by Lemma 2.3, which may be applied thanks to (7.64), (3.59), (3.61) and ε small enough.

Finally (7.66) for $\mathcal{S}^{-1}$ follows by

$$
\partial_u \mathcal{S}^{-1}(u)[h] = -\mathcal{S}^{-1}(u)\,\partial_u \mathcal{S}(u)[h]\,\mathcal{S}^{-1}(u)\,,
$$

and (2.7), (7.64) for $\mathcal{S}^{-1}$, and (7.66) for $\mathcal{S}$.

20. — The operatpr $\mathcal{R}$, defined in (3.55) , where r_0, r_{-1} are defined in (3.52), (3.53), satisfies the following estimates:

$$
\begin{align}
\left|\mathcal{R}\right|_s &\leq \varepsilon C(s)(1 + \|u\|_{s+2\tau_0+2\mathfrak{s}_0+12}) \tag{7.67} \\
\left|\mathcal{R}\right|_s^{\mathrm{Lip}(\gamma)} &\leq \varepsilon C(s)(1 + \|u\|_{s+2\tau_0+2\mathfrak{s}_0+13}^{\mathrm{Lip}(\gamma)}) \tag{7.68} \\
\left|\partial_u \mathcal{R}(u)[h]\right|_s &\leq \varepsilon C(s)\big(\|h\|_{s+2\tau_0+2\mathfrak{s}_0+12} + \|u\|_{s+2\tau_0+2\mathfrak{s}_0+13}\|h\|_{2\tau_0+3\mathfrak{s}_0+12}\big). \tag{7.69}
\end{align}
$$

Let $T := r_0 + r_{-1}\partial_x^{-1}$. By (2.5), (2.6), (6.5), (7.61), (7.62), (7.58), (7.59), (7.51), (7.22), and using the trivial fact that $|\partial_x^{-1}|_s \leq 1$ and $|\pi_0|_s \leq 1$ for all $s \geq 0$, we get

$$
\begin{align}
\left|T\right|_s &\leq \varepsilon C(s)(1 + \|u\|_{s+2\tau_0+2\mathfrak{s}_0+12}) \tag{7.70} \\
\left|T\right|_s^{\mathrm{Lip}(\gamma)} &\leq \varepsilon C(s)(1 + \|u\|_{s+2\tau_0+2\mathfrak{s}_0+13}^{\mathrm{Lip}(\gamma)}). \tag{7.71}
\end{align}
$$

Differentiating T with respect to u, and using (2.5), (6.5), (7.63), (7.60), (7.51), (7.22) and (7.23), one has

$$
\left|\partial_u T(u)[h]\right|_s \leq \varepsilon C(s)\big(\|h\|_{s+2\tau_0+2\mathfrak{s}_0+12} + \|u\|_{s+2\tau_0+2\mathfrak{s}_0+13}\|h\|_{2\tau_0+3\mathfrak{s}_0+12}\big). \tag{7.72}
$$

Finally (2.7), (2.10) (7.64)-(7.66), (7.70)-(7.72) imply the estimates (7.67)-(7.69).

21. — Using Lemma 6.5, (3.59) and all the previous estimates on $\mathcal{A}, B, \rho, \mathcal{M}, \mathcal{T}, \mathcal{S}$, the operators $\Phi_1 = \mathcal{A}B\rho\mathcal{M}\mathcal{T}\mathcal{S}$ and $\Phi_2 = \mathcal{A}B\mathcal{M}\mathcal{T}\mathcal{S}$, defined in (3.57), satisfy (3.60) (note that $\sigma > 2\tau_0 + 2\mathfrak{s}_0 + 9$). Finally, if the condition (3.61) holds, we get the estimate (3.62).

The other estimates (3.63)-(3.68) follow by (7.22), (7.23), (7.51), (7.67)-(7.69). The proof of Lemma 3.2 is complete.

Proof of Lemma 3.3. For each fixed $\varphi \in \mathbb{T}^\nu$, $\mathcal{A}(\varphi)h(x) := h(x + \beta(\varphi, x))$. Apply (6.16) to the change of variable $\mathbb{T} \to \mathbb{T}$, $x \mapsto x + \beta(\varphi, x)$:

$$
\|\mathcal{A}(\varphi)h\|_{H_x^s} \leq C(s)\big(\|h\|_{H_x^s} + |\beta(\varphi, \cdot)|_{W^{s,\infty}(\mathbb{T})}\|h\|_{H_x^1}\big).
$$

Since $|\beta(\varphi, \cdot)|_{W^{s,\infty}(\mathbb{T})} \leq |\beta|_{s,\infty}$ for all $\varphi \in \mathbb{T}^\nu$, by (7.5) we deduce (3.69). Using (6.17), (3.59), and (7.5),

$$
\|(\mathcal{A}(\varphi) - I)h\|_{H_x^s} \lesssim_s |\beta|_{L^\infty}\|h\|_{H_x^{s+1}} + |\beta|_{s,\infty}\|h\|_{H_x^2} \lesssim_s \varepsilon\big(\|h\|_{H_x^{s+1}} + \|u\|_{s+\mathfrak{s}_0+3}\|h\|_{H_x^2}\big).
$$

By (7.8), estimates (3.69) and (3.70) also hold for $\mathcal{A}(\varphi)^{-1} = \mathcal{A}^{-1}(\varphi) : h(y) \mapsto h(y + \tilde{\beta}(\varphi, y))$.

The multiplication operator $\mathcal{M}(\varphi) : H_x^s \to H_x^s$, $\mathcal{M}(\varphi)h := v(\varphi, \cdot)h$ satisfies

$$
\begin{align}
\|(\mathcal{M}(\varphi) - I)h\|_{H_x^s} = \|(v(\varphi, \cdot) - 1)h\|_{H_x^s} &\lesssim_s \|v(\varphi, \cdot) - 1\|_{H_x^s}\|h\|_{H_x^1} + \|v(\varphi, \cdot) - 1\|_{H_x^1}\|h\|_{H_x^s} \\
&\lesssim_s \|v - 1\|_{s+\mathfrak{s}_0}\|h\|_{H_x^1} + \|v - 1\|_{1+\mathfrak{s}_0}\|h\|_{H_x^s} \lesssim_s \varepsilon\big(\|h\|_{H_x^s} + \|u\|_{s+\tau_0+2\mathfrak{s}_0+6}\|h\|_{H_x^1}\big) \tag{7.73}
\end{align}
$$

by (6.5), (2.5), Lemma 2.4, (7.42) and (3.59). The same estimate also holds for $\mathcal{M}(\varphi)^{-1} = \mathcal{M}^{-1}(\varphi)$, which is the multiplication operator by $v^{-1}(\varphi, \cdot)$. The operators $\mathcal{T}^{\pm 1}(\varphi) h(x) = h(x \pm p(\varphi))$ satisfy

$$\|\mathcal{T}^{\pm 1}(\varphi) h\|_{H_x^s} = \|h\|_{H_x^s}, \quad \|(\mathcal{T}^{\pm 1}(\varphi) - I) h\|_{H_x^s} \leq \varepsilon \gamma_0^{-1} C \|h\|_{H_x^{s+1}}, \tag{7.74}$$

by (6.17), (3.59), (7.52) and by the fact that $p(\varphi)$ is independent on the space variable.

By (2.12), (7.64), (3.59) and Lemma 2.4, the operator $\mathcal{S}(\varphi) = I + w(\varphi, \cdot) \partial_x^{-1}$ and its inverse satisfy

$$\|(\mathcal{S}^{\pm 1}(\varphi) - I) h\|_{H_x^s} \leq_s \varepsilon \big(\|h\|_{H_x^s} + \|u\|_{s+2\tau_0+3\mathfrak{s}_0+9} \|h\|_{H_x^1} \big). \tag{7.75}$$

Collecting estimates (7.73), (7.74), (7.75) we get (3.71) and (3.72). Lemma 3.3 is proved.

References

[1] Baldi P., *Periodic solutions of forced Kirchhoff equations*, Ann. Scuola Norm. Sup. Pisa, Cl. Sci. (5), Vol. 8 (2009), 117-141.

[2] Baldi P., *Periodic solutions of fully nonlinear autonomous equations of Benjamin-Ono type*, to appear on Ann. I. H. Poincaré (C) Anal. Non Linéaire.

[3] Bambusi D., Graffi S., *Time quasi-periodic unbounded perturbations of Schrödinger operators and KAM methods*, Commun. Math. Phys. 219 (2001), 465-480.

[4] Berti M., *Nonlinear oscillations of Hamiltonian PDEs*, Progress in Nonlin. Diff. Eq. Appl., 74, Birkhäuser, Boston, 1-181, 2008.

[5] Berti M., Biasco L., *Branching of Cantor manifolds of elliptic tori and applications to PDEs*, Comm. Math. Phys, 305, 3, 741-796, 2011.

[6] Berti M., Bolle P., *Cantor families of periodic solutions for completely resonant nonlinear wave equations*, Duke Math. J. 134 (2006) 359-419.

[7] Berti M., Bolle P., *Cantor families of periodic solutions of wave equations with C^k nonlinearities*, NoDEA 15 (2008), 247-276.

[8] Berti M., Bolle P., *Quasi-periodic solutions with Sobolev regularity of NLS on $\mathbb{T}^d$ with a multiplicative potential*, Eur. Jour. Math, Vol. 15, 229-286, 2013.

[9] Berti M., Bolle P. *Sobolev quasi periodic solutions of multidimensional wave equations with a multiplicative potential*, Nonlinearity, 25, 2579-2613, 2012.

[10] Berti M., Bolle P., Procesi M., *An abstract Nash-Moser theorem with parameters and applications to PDEs*, Ann. I. H. Poincaré, 27, 377-399, 2010.

[11] Berti M., Biasco P., Procesi M., *KAM theory for the Hamiltonian DNLW*, to appear on Annales Scientifiques de l'ENS.

[12] Berti M., Biasco P., Procesi M., *KAM theory for the reversible derivative wave equation*, preprint 2012.

[13] Bourgain J., *Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE*, Internat. Math. Res. Notices, no. 11, 1994.

[14] Bourgain J., *Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations*, Annals of Math. 148, 363-439, 1998.

[15] Bourgain J., *Periodic solutions of nonlinear wave equations*, Harmonic analysis and partial differential equations, Chicago Lectures in Math., Univ. Chicago Press, (1999), pp.69-97.

[16] Bourgain J., *Green's function estimates for lattice Schrödinger operators and applications*, Annals of Mathematics Studies 158, Princeton University Press, Princeton, 2005.

[17] Craig W., *Problèmes de petits diviseurs dans les équations aux dérivées partielles*, Panoramas et Synthèses, 9, Société Mathématique de France, Paris, 2000.

[18] Craig W., Wayne C. E., *Newton's method and periodic solutions of nonlinear wave equation*, Comm. Pure Appl. Math. 46, 1409-1498, 1993.

[19] Delort J.-M., *A quasi-linear Birkhoff normal forms method. Application to the quasi-linear Klein-Gordon equation on* $\mathbb{S}^1$, Astérisque 341 (2012).

[20] Eliasson L.H., Kuksin S., *KAM for non-linear Schrödinger equation*, Annals of Math., 172, 371-435, 2010.

[21] Eliasson L. H., Kuksin S., *On reducibility of Schrödinger equations with quasiperiodic in time potentials*, Comm. Math. Phys, 286, 125-135, 2009.

[22] Geng J., Xu X., You J., *An infinite dimensional KAM theorem and its application to the two dimensional cubic Schrödinger equation*, Adv. Math. 226, 6, 5361-5402, 2011.

[23] Gentile, G., Procesi, M., *Periodic solutions for a class of nonlinear partial differential equations in higher dimension*, Comm. Math. Phys. 289 (2009), no. 3, 863-906, 2009.

[24] Grebert B., Thomann L., *KAM for the quantum harmonic oscillator*, Comm. Math. Phys. 307, 2, 383-427, 2011.

[25] Hamilton R.S., *The inverse function theorem of Nash and Moser*, Bull. Amer. Math. Soc. 7 (1) (1982), 65-222.

[26] Hörmander L., *The boundary problems of physical geodesy*, Arch. Ration. Mech. Anal. 62 (1976), no. 1, 1-52.

[27] Iooss G. , Plotnikov P.I., Toland J.F., *Standing waves on an infinitely deep perfect fluid under gravity*, Arch. Ration. Mech. Anal. 177 (2005), no. 3, 367-478.

[28] Iooss G. , Plotnikov P.I., *Small divisor problem in the theory of three-dimensional water gravity waves*, Mem. Amer. Math. Soc. 200, (2009), no. 940.

[29] Iooss G. , Plotnikov P.I., *Asymmetrical three-dimensional travelling gravity waves*, Arch. Ration. Mech. Anal. 200 (2011), no. 3, 789880.

[30] Kappeler T., Pöschel J., *KAM and KdV*, Springer, 2003.

[31] Klainermann S., Majda A., *Formation of singularities for wave equations including the nonlinear vibrating string*, Comm. Pure Appl. Math., 33, 241-263, 1980.

[32] Kuksin S., *Hamiltonian perturbations of infinite-dimensional linear systems with imaginary spectrum*, Funktsional. Anal. i Prilozhen. 21, no. 3, 22–37, 95, 1987.

[33] Kuksin S., *A KAM theorem for equations of the Korteweg-de Vries type*, Rev. Math-Math Phys., 10, 3, 1-64, 1998.

[34] Kuksin S., *Analysis of Hamiltonian PDEs*, Oxford Lecture Series in Mathematics and its Applications, 19. Oxford University Press, 2000.

[35] Kuksin S., Pöschel J., *Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation*, Annals of Math. (2) 143, 149-179, 1996.

[36] Lax P., *Development of singularities of solutions of nonlinear hyperbolic partial differential equations*, Journal of Mathematical Physics, 5, 5, 1964.

[37] Liu J., Yuan X., *A KAM Theorem for Hamiltonian Partial Differential Equations with Unbounded Perturbations*, Comm. Math. Phys, 307 (3), 629-673, 2011.

[38] Moser J., *A rapidly convergent iteration method and non-linear partial differential equations - I*, Ann. Sc. Norm. Sup. Pisa Cl. Sci. III Ser., Vol. 20, no. 2, (1966), page 265-315.

[39] Pöschel J., *A KAM-Theorem for some nonlinear partial differential equations*, Ann. Scuola Norm. Sup. Pisa Cl. Sci.(4), 23, 119-148, 1996.

[40] Procesi C., Procesi M., *A KAM algorithm for the completely resonant nonlinear Schrödinger equation*, preprint 2012.

[41] Procesi M., Xu X., *Quasi-Töplitz Functions in KAM Theorem*, preprint 2011.

[42] Rabinowitz P.H., *Periodic solutions of nonlinear hyperbolic partial differential equations*, Part I and II, Comm. Pure Appl. Math., Vol. 20 (1967), 145–205, and Vol. 22 (1969), 15-39.

[43] Wang W. M., *Supercritical nonlinear Schrödinger equations I: quasi-periodic solutions*, preprint arxiv.org.

[44] Wayne E., *Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory*, Comm. Math. Phys. 127, 479-528, 1990.

[45] Zhang J., Gao M., Yuan X. *KAM tori for reversible partial differential equations*, Nonlinearity 24, 1189-1228, 2011.

Massimiliano Berti, Pietro Baldi, Dipartimento di Matematica e Applicazioni "R. Caccioppoli", Università degli Studi Napoli Federico II, Via Cintia, Monte S. Angelo, I-80126, Napoli, Italy, `m.berti@unina.it`, `pietro.baldi@unina.it`.

Riccardo Montalto, SISSA, Via Bonomea 265, 34136, Trieste, Italy, `riccardo.montalto@sissa.it`.